# Global governance as promise-making.

## Negotiating and monitoring learning goals in the time of SDGs

Author: Clara Fontdevila

Supervisor: Prof. Antoni Verger

February 2021

# Acknowledgments

First and above, I would like to express my deepest gratitude for my supervisor, Toni Verger, for his generosity and encouragement. This thesis owns much to his thoughtful comments. Thanks for all the opportunities to learn and for the support along the way – it made all the difference.

My heartfelt thanks go also to all the members of the GEPS team – for the advice and kindness, for all the stimulating discussions. They have been a constant source of inspiration over these years, and their companionship has made this journey a much more enjoyable and meaningful one.

I am indebted to the many people who found the time to be interviewed and agreed to share their experiences and perspectives. While I cannot acknowledge them individually, I do wish to express my sincere gratitude, as this thesis would not have been possible without their kind collaboration.

I am also thankful to Prof. Oren Pizmony-Levy, who welcomed me at Teachers College; and to Prof. Kerstin Martens, who welcomed me the University of Bremen. These were formative experiences that very much helped me to shape this thesis.

My gratitude goes also to the members and collaborators of UNESCO's ERF team, for their hospitality and trust – it has been a real privilege. I wish especially to express my gratitude to Maya Prince, for her infinite wisdom and kindness, and for her patience with all my questions and indecision.

Finally, thanks to my friends and family, from Manresa and elsewhere – for their unrelenting encouragement and moral support, and for bringing things into perspective. Many thanks in particular to my parents and J., for their understanding and for being always there. Moltes gràcies.

---

# Table of contents

# List of tables and figures

# List of abbreviations and acronyms

| | |
|---|---|
| AAAA | Addis Ababa Action Agenda |
| ACER | Australian Council for Educational Research |
| ACER-GEM | ACER - Centre for Global Education Monitoring |
| ADG-E | Assistant Director-General for Education |
| ALO | Assessment of Learning Outcomes |
| ANCEFA | Africa Network Campaign on Education for All |
| ASER | Annual Status of Education Report |
| ASPBAE | Asia South Pacific Association for Basic and Adult Education |
| BESTA | Better Education Statistics and global Action to improve learning |
| BICSE | Board of International Comparative Studies in Education |
| CCNGO/EFA | Collective Consultation of NGOs on Education for All |
| CGD | Center for Global Development |
| CNA | Cross-National Assessment |
| CNL | Children Not Learning |
| CONFEMEN | *Conférence des ministres de l'Éducation des États et gouvernements de la Francophonie* [Conference of the Ministers of Education of French speaking countries] |
| CSO | Civil Society Organization |
| CUE | Center for Universal Education |
| DAC | Development Assistance Committee (OECD) |
| DEC | Development Economics Vice Presidency (World Bank) |
| DFAT | Australian Department of Foreign Affairs and Trade |
| DfID | Department for International Development |
| ECD | Early Childhood Development |
| ECOSOC | United Nations Economic and Social Council |
| EFA | Education for All |
| EGMA | Early Grade Mathematics Assessment |
| EGRA | Early Grade Reading Assessment |
| EI | Education International |
| ESRC | Economic and Social Research Council |
| ETAG | Extended Technical Advisory Group |
| EU | European Union |
| FFA | Framework for Action |
| FfD | Financing for Development |

| | |
|---|---|
| FTI | Fast Track Initiative |
| FTI-SC | Fast Track Initiative – Steering Committee |
| GAD | Gender and Development |
| GAML | Global Alliance to Monitor Learning |
| GCE | Global Campaign for Education |
| GCL | Global Compact on Learning |
| GCLO | Global Comparability of Learning Outcomes |
| GDP | Gross Domestic Product |
| GEFI | Global Education First Initiative |
| GEM | Global Education Meeting |
| GEMR | Global Education Monitoring Report |
| GLAD | Global Learning Assessment Database |
| GMR | Global Monitoring Report |
| GPE | Global Partnership for Education |
| HCI | Human Capital Index |
| HCP | Human Capital Project |
| HDI | Human Development Index |
| HLF | High-Level Forum (EFA) |
| HLG | High-Level Group (EFA) |
| HLG-PCCB | High-level Group for Partnership, Coordination and Capacity-Building for statistics for the 2030 Agenda for Sustainable Development |
| HLO | Harmonized Learning Outcomes |
| HLP | High-Level Panel of Eminent Persons |
| HLPF | High-Level Political Forum |
| IAC | Inter-Agency Commission |
| IAEG-SDGs | Inter-Agency and Expert Group on SDG Indicators |
| IBE | International Bureau of Education |
| ICAE | International Council for Adult Education |
| ICESDF | Intergovernmental Committee of Experts on Sustainable Development Finance |
| IDG | International Development Goal |
| IEA | International Association for the Evaluation of Educational Achievement |
| IGN | Intergovernmental Negotiations |
| IIEP | International Institute for Educational Planning - UNESCO |

| | |
|---|---|
| ILO | International Labour Organization |
| IMF | International Monetary Fund |
| INGO | International Non-Governmental Organization |
| IO | International Organization |
| IOR | Inter-Organizational Relations |
| IR | International Relations |
| LAMP | Literacy Assessment and Monitoring Programme |
| LaNA | Literacy and Numeracy Assessment |
| LAYS | Learning-Adjusted Years of Schooling |
| LeAP | Learning Assessment Platform |
| LLECE | *Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación* [Latin American Laboratory for Assessment of the Quality of Education] |
| LMP | Learning Metrics Partnership |
| LMTF | Learning Metrics Task Force |
| LPT | Learning Poverty Target |
| LSLA | Large-Scale Learning Assessment |
| LT | Learning Poverty Target |
| MD | Millennium Declaration |
| MDG | Millennium Declaration Goal |
| MoU | Memorandum of Understanding |
| MSI | Management Systems International |
| NA | National Assessment |
| NGO | Non-Governmental Organization |
| OECD | Organisation for Economic Co-operation and Development |
| OLO | Observatory of Learning Outcomes |
| OWG | Open Working Group |
| PASEC | *Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN* [Programme for the Analysis of Education Systems] |
| PILNA | Pacific Islands Literacy and Numeracy Assessment |
| PISA | Programme for International Student Assessment |
| PISA-D | PISA for Development |
| PIRLS | Progress in International Reading Literacy Study |
| RBM | Results-Based Management |
| READ | Russia Education Aid for Development |
| SABER | Systems Approach for Better Education Results |

| | |
|---|---|
| SABER-UF | SABER Umbrella Facility |
| SACMEQ | Southern and Eastern Africa Consortium for Monitoring Educational Quality |
| SDG | Sustainable Development Goal |
| SDSN | Sustainable Development Solutions Network |
| SEA-PLM | Southeast Asia Primary Learning Metrics |
| SPC | Strategic Planning Committee |
| TAG | Technical Advisory Group |
| TCG | Technical Cooperation Group |
| TIMSS | Trends in International Mathematics and Science Study |
| TST | Technical Support Team |
| TVET | Technical and Vocational Education and Training |
| UIS | UNESCO Institute for Statistics |
| UIS-RS | UIS Universal Reporting Scales |
| UN | United Nations |
| UN-SG | United Nations Secretary-General |
| UN/DESA | United Nations Department of Economic and Social Affairs |
| UNDG | United Nations Development Group |
| UNESCO | United Nations Educational, Scientific and Cultural Organization |
| UNFPA | United Nations Population Fund |
| UNGA | United Nations General Assembly |
| UNHCR | United Nations High Commissioner for Refugees |
| UNICEF | United Nations International Children's Emergency Fund |
| UNPD | United Nations Development Programme |
| UNSC | United Nations Statistical Commission |
| UNSD | United Nations Statistical Division |
| UNTT | UN System Task Team on the post-2015 UN Development Agenda |
| UPE | Universal Primary Education |
| USAID | United States Agency for International Development |
| WB | The World Bank |
| WCEFA | World Conference on Education for All |
| WEF | World Education Forum |
| WEI | World Education Indicators |
| WER | World Education Report |
| WHO | World Health Organization |

# Chapter 1

## Introduction

### 1.1. Background of the research

Goal- and target-setting have a long history in the field of education. While the two education-related Millennium Development Goals and the six Education For All goals feature among the most well-known antecedents of the Sustainable Development Goal 4-Education 2030 Agenda, the practice of goal-setting in education dates back at least as far as the 1960s, with the adoption of different sets of education goals and time-bound targets in the context of UNESCO's regional conferences (King, 2016). In fact, calls for the realization of the right to education can be traced back to the creation of UNESCO and the Universal Declaration of Human Rights adopted in 1948. Since then, pronouncements and public commitments to the universal access to education have been reiterated countless times in a variety of international conferences and declarations (Chabbott, 2003).

While the education goals adopted during the 1960s represent the first instance of goal-setting in the context of the UN system, such practices expanded rapidly within the international development realm. Since the 1960s, a wide array of aspirational goals and quantitative targets have been adopted in a variety of areas, frequently as an outcome of UN-sponsored conferences and summits (Fukuda-Parr, 2014; Jolly, 2004; Hulme, 2009), to the extent that today, goal-setting is regarded as a UN legacy, and constitutes one of its procedural hallmarks (Jolly, Emmerji and Weiss, 2009). However, goal-setting practices reached an unprecedented level of attention and public exposure with the adoption of the Millennium Declaration and the subsequent approval of the Millennium Development Goals (MDGs) in the early 2000s. Despite being modeled on the International Development Goals set out by the OECD's DAC in 1996 (Black and White, 2003), and in spite of their uneven record of achievement (Hulme and Scott, 2010), the MDGs have frequently been touted as a turning point – not only given their success as an advocacy tool and in mobilizing political support, but also as they represented the most comprehensive set of internationally-agreed goals and targets adopted to date (Fukuda-Parr and Yamin, 2013; Waage et al., 2010).

The adoption of the Sustainable Development Goals (SDGs) in the UN General Assembly of 2015 adds to a well-established and relatively institutionalized tradition within the development realm, and within the UN system in particular. In this sense, the SDGs represent a form of continuity in that they rely on mediums and procedures that today constitute routinized practice for many development stakeholders. At the same time, the SDGs represent a departure from MDGs – not only in their content, but also in procedural terms, that is, in the political nature of the process behind them. In relation to the former, the new set of goals has entailed a normative shift – expanding the very notion of

development beyond the focus on poverty and its conceptualization as a North-South endeavor that characterized the MDGs. This has ensured the universal character of the agenda, as well as a clearer focus on the reduction of inequality (Biermann, Kanie and Kim, 2017). In relation to the latter, the openly-negotiated nature of SDGs contrasts with the technocratic character of MDGs. Different authors have emphasized the consultative and negotiated character of SDGs – drawing attention to the role played by the Open Working Group (OWG) of the UN General Assembly in complementing the work of the High-Level Task Force of Eminent Persons (HLP) directly appointed by the UN Secretary-General (Dodds et al., 2017; Fukuda-Parr and McNeill, 2019). As an inter-governmental group having benefited from the input of different constituencies, the OWG would have enabled the participation of a wider range of stakeholders, thus ensuring the consensual, open and inclusive character of the negotiation.

In addition to these normative and political shifts, Fukuda-Parr and McNeill (2019) have argued that the SDGs entailed a methodological shift, moving development agendas towards the governance-by-numbers paradigm (cf. Ozga, 2009; Rose, 1999). While the use of quantitative targets is not new, the specificity and novelty of SDGs lies in the fact that, for the first time, development goals were not identified after the adoption of the agenda (i.e., the public declaration of political commitment establishing certain priorities). Rather, the final declaration (i.e., the public statement) was based directly on the previously agreed goals and targets. To put it differently, goal-setting practices were not a mere (ex-post) operativization effort, but explicitly conceived as the core of the new agenda. Thus, their centrality occurred less by accident (as in the case of MDGs), and more by design.

Finally, it is worth noting that, in the case of education, the adoption of SDG4 (i.e., the "education goal", also known as the Education 2030 Agenda) represents a major breakthrough in that the new agenda has involved the convergence of the two sets of goals previously in place i.e., the two education-related MDGs and the six Education For All (EFA) goals (King, 2017; Unterhalter, 2019). Another relevant education-specific change brought about by SDG4 is the growing relevance of learning-related indicators for the monitoring procedures established by SDG4. Indeed, one of the most significant shifts entailed by the new agenda is the focus on learning outcomes – a transformation sometimes referred to as the quality turn, and defined as an effort to transcend a focus on schooling and enrolment figures as key indicators or progress (Sachs-Israel, 2017; Sayed, Ahmed and Mogliacci, 2018). Thus, large-scale learning assessments feature prominently in the indicator framework associated with the SDG4 agenda as key data sources to monitor progress on learning. Relatedly, both the global and the thematic indicator frameworks established for the monitoring of Education 2030 are unambiguous on the need for countries to adopt or participate in some form of learning assessment so that student achievement can be reported using an internationally comparable scale.

## 1.2. Defining the research puzzle

Given their (self-professed) groundbreaking character, the SDGs have sparked extensive debate and interest within development circles. However, both public comment and scholarship have tended to focus on the (potential) effectiveness and impact of the new agenda – or lack thereof. Celebratory accounts coexist with more skeptical reflections. Whilst the SDGs have been officially heralded as a revolutionary turn (Dodds et al., 2017), other observers of the process have adopted a more critical approach, pointing to the lack of effective accountability mechanisms and the limited actionability of the new goals – see for instance Easterly (2015). In an in-between position, many approach the new goals as a promising initiative while simultaneously drawing attention to the implementation difficulties of such an overarching framework (Biermann et al., 2017)[1]. Similar debates (pointing to analogous shortcomings) have been held in the area of education – for an overview, see Unterhalter, Poole, and Winters (2015).

In contrast to the profusion of evaluative judgments and comments, empirical work on the crafting and negotiation of the new agenda remains limited. Existing accounts of the negotiation process have generally adopted a bird's-eye view, focusing on the (formal) architecture of the consultation and decision-making processes, and examining the new set of goals and targets *en bloc*, that is, addressing SDGs as a whole rather than delving into sector-specific dynamics (see for instance Brenner, 2015; Dodds, Donoghue and Roesch, 2017; and Kamau, Chasek and O'Connor, 2018). While fine-grained and informative in their own right, such accounts tend to neglect or overlook "soft" policy areas (including education) – and their focus of interest usually gravitates to those areas characterized by high politics dynamics. These accounts have thus tended to adopt a state-centric approach – precisely because state representatives tend to be more active, familiar and opinionated in "hard policy areas" that have long been the subject of national diplomacies.

In the case of education, available analysis of the formulation of SDG4 goal, targets and indicators are equally scarce. It is true that recent accounts inquire into the process of negotiation, highlighting contentious issues and mapping the wide range of processes that eventually gave rise to (or galvanized into) the current wording of the goal and the associated targets and metrics (see for instance King, 2017; Sachs-Israel, 2017; Sayed and Ahmed, 2015; or Unterhalter, 2019). However, since such questions generally do not constitute the main focus of these works, they are addressed in a necessarily superficial and succinct way, which tends to overlook the impact of intra-organizational dynamics or the role of more informal negotiations.

Adding to its paucity, a second limitation of the literature on the negotiation of SDGs (or SDG4) lies in the fact that, while insightful, most of the available works follow an

---

[1] Such positions and debates have tended to mirror those arisen by the MDGs – for an overview, see Hulme (2009).

essentially descriptive approach and do not engage in a dialogue with the main contributions made by the scholarship on global governance, and/or International Organizations (IOs). In fact, some of the most comprehensive efforts in this area adopt the form of inside accounts – see for instance Dodds, Donoghue and Roesch, 2017; or Kamau, Chasek and O'Connor, 2018.

While this is not inherently problematic, the lack of cross-fertilization between scholarship on international governance and the available overviews on the negotiation of SDGs represents a missed opportunity to gain a better understanding of the micro-foundations of transnational norm-building dynamics to which goal- setting and target-setting practices contribute. This is all the more relevant given that, while the constructivist agenda in international relations (and the study of IOs in particular) has long emphasized the key role of norms, there is limited theoretical work on the very *emergence* and *codification* of these norms within IOs. Indeed, the uneven attention given to the different stages of the life cycle of global norms was noted by Finnemore and Sikkink more than two decades ago, right after the ideational turn of the late 1980s that secured a return to normative concerns (Finnemore and Sikkink, 1998). The authors observed that research on international norms tended to concentrate on the later stages of their life cycle – that is, norm cascade (or contagion) and norm internalization (their consolidation and universalization). The very creation of these norms, however, would have remained much more of a blind spot.

Although significant inroads have been made over recent years into an understanding of this norm-making labor, some theoretical gaps persist. First, as noted by Kentikelenis and Seabrooke (2017), while there is an increasingly clear understanding of the role of inter-organizational dynamics in the dissemination and institutionalization of global norms, the intra-organizational development of these norms remains poorly understood. In other words, it remains unclear how certain norms are produced and institutionalized as desirable *within* a given organization. Second, while global norms are articulated by IOs in a wide array of texts and legal technologies, (i.e., mediums and scripts in which diagnosis and prescriptions crystallize) (Halliday, Block-Lieb and Carruthers, 2009), International Relations (IR) literature has tended to focus on a narrow set of them. Those articulated as programmatic claims and statements of shared aspirations, but devoid of resources for their implementation and not linked to a clear set of policy rules (as would be the case of SDGs), feature among the less systematically scrutinized. Given that in the field of development the UN has specialized on these programmatic products oriented towards setting priorities and moral standards, and appealing to shared principles and values (Babb and Chorev, 2016; Barnett and Finnemore, 2018), the limited theorization of such practices remains problematic for an understanding of the role and influence of the UN system from a constructivist perspective. Third, as noted by Park (2005), existing accounts of an IO's norm creation have not succeeded in explaining how normative shifts occur within a given organization, or how an IO's norms of choice are affected by the priorities espoused by other institutions, and non-state actors in particular. Overall, these

empirical and theoretical gaps suggest that, as noted by Wiener (2003), the emergence (and waning) of norms remains a crucial challenge within IR scholarship.

Finally, a third question that the literature on the SDGs (and other normative frameworks) has tended to overlook is that of the effects of the new agenda on the organizational practices of the partaking IOs. As noted above, there are extensive comments on the extent to and ways through which the new goals are likely to shape the behavior of other actors (particularly states and donors). There is, however, a more limited understanding of how SDGs impact the practices of IOs engaged in such efforts, in particular, their organizational culture, i.e., the ideologies, worldviews, incentives, norms and routines shaping their behavior. While such questions are touched upon in most of the reflections on the impact of SDGs, they have not received systematic consideration. This represents a problematic absence in that, although the SDGs reproduce the state-centric approach to development accountability (Sexsmith and McMichael, 2015), their ultimate impact is at least partially mediated by the practices of IOs[2]. Yet, there is limited evidence on how (and to what extent) the institutional agendas and behavior of IOs are permeated by the priorities encapsulated in SDGs.

The limited discussion on the impact of the quantification needs of the new agenda over the practices of concerned IOs is particularly indicative of this neglect. Certainly, the generalization of global rankings, benchmarking and target-setting practices over the last few decades has translated into a proliferation of studies on the cognitive and regulatory effects of such measurement efforts – see for instance Broome and Quirk, 2015a; Fukuda-Parr and Yamin, 2013; or Merry, 2011. However, there is a more limited understanding of the impact of this quantification imperative on the mundane practices of IOs in charge of their production or harmonization. In the context of the SDGs, such an omission is particularly problematic in that the measurement and monitoring instruments consolidated by the new framework have major implications and directly affect the normative cycles and data-collection priorities of IOs and other transnational actors.


## 1.3. Objective of the study and research questions

On the basis of the considerations described above, this dissertation argues that the negotiation and measurement of SDG4 can be productively analyzed from a constructivist perspective, and that such an endeavor is of theoretical significance. This is because, whilst there is no dearth of literature addressing the potential and impact of SDGs (and of SDG4 in particular), other aspects of the new agenda remain comparatively under-researched and under-theorized – especially those that conventionally fall within the constructivist research agenda. The dissertation departs from the assumption that there is limited understanding of the interplay between the instruments (goals, targets and indicators) put forward by the new agenda, and the organizational cultures of the wide

---

[2] For a similar argument on the MDGs, see Hulme and Scott (2010).

range of IOs with a mandate in the global governance of education (and consequently, involved in the crafting, monitoring and measurement of the new agenda). More specifically, it remains unclear how these organizations shape the construction of these distinctly prescriptive and aspirational norms, and how their own functioning and priorities are shaped by these normative frameworks.

I also argue that the relevance of these questions goes beyond the SDG4 agenda – it may also contribute to a better understanding of the practice of global goal-setting as a key component of transnational norm-building. Despite constituting a well-established and institutionalized practice within the development realm, the production of prescriptive and aspirational goals remains an under-theorized practice. Consequently, the empirical examination of SDG4 constitutes a fruitful opportunity to gain insight into the sociological micro-foundations of such practices.

*The main goal of this dissertation is to examine the negotiation and monitoring of Sustainable Development Goal 4 by bringing to the center of the analysis the culture and interests of the international organizations partaking in these endeavors, as well as the institutional architecture in which inter-organizational interactions took place.*

This general objective, in turn bifurcates into two separate sets of research goals and questions that warrant further investigation, namely:

- *Research objective 1 – The practice of promise-making*

  The first research goal of the dissertation is to understand how the normative statements and associated targets encapsulated in SDG4 are agreed upon and negotiated within and between those international organizations with a mandate or agenda on the global governance of education.

  The specific objectives included in this line of inquiry are as follows:

  1.1 Reconstruct and analyze the sequence of inter-governmental, and intra- and inter-organization negotiations spanning from mid-2012 to the end of 2015, which culminated in the adoption of a common set of goals and targets[3].

  1.2 Identify and map the constellation of collective and individual actors involved in the process, and reconstruct the rationales driving their participation in the process, as well as their priorities and modalities of engagement.

  1.3 Identify the varying degrees of influence exerted by different organizations, as well as the key *locus* of decision-making.

  1.4 Examine the main focus of debate, contention and (open or covert) disagreement within and between the involved organizations.

---

[3] That is, the run-up to the World Education Forum 2015 and to the final adoption of the Education 2030 Framework for Action in November 2015.

1.5 Identify key mechanisms of consensus-building operating within and across the partnering organizations.

- *Research objective 2 – The practice of promise-tracking*

The second research goal of the dissertation is to understand how the quantification needs brought about by the new agenda have impacted on the institutional agendas of the organizations involved in the production of SDG4 indicators, as well as on the relationships between data suppliers and data harmonizers – specifically, in the production of learning metrics and the monitoring of learning outcomes.

The specific objectives addressed by this line of inquiry are as follows:

2.1. Analyze the ways in which the impetus given to learning metrics in the SDG4 agenda has penetrated the institutional agendas and priorities of the main organizations with a role or mandate in the production and harmonization of assessment data.

2.2. Examine how and to what extent the commensuration and coordination needs brought about by the new monitoring mechanisms have affected the relations of political, technical and symbolic interdependence among the producers and the harmonizers of learning data.

2.3. Examine how the new quantification needs have transformed the hierarchization dynamics governing the assessment field (that is, the distribution of different forms of authority).

2.4. Identify the strategies mobilized by these organizations in order to preserve or access dominant or influential positions, including emerging and/or changing patterns of cooperation and competition.

## 1.4. Empirical strategy

In order to address the research questions outlined above, this dissertation relies on a combination of three data-collection methods. The first and main data source is a corpus of semi-structured interviews conducted with a variety of stakeholders with a significant involvement in the negotiation, monitoring and reporting efforts that constitute the subject of the dissertation. Recruited through a combination of purposive and snowballing sampling strategies, most of these informants fall within the *elite interviewing* category. This is an issue with key implications for access strategies, the design of the interview protocol, and data analysis (Harvey, 2011). It is of particular relevance given the need to reflect and make sense of the *interpretation* given by the interviewed actors of the processes or events they report (that is, going beyond descriptive accounts) as required by a constructivist research agenda.

Complementarily, the research also relies on a large body of documentary data from the key documents produced throughout the post-2015 process – including memos and position statements issued by involved organizations, political *communiqués*, technical reports and policy briefs. These documents allowed to complement (and add nuance to) the insights gained through interview and observation. Given the challenges inherent in "studying up" (Nader, 1972), such strategy was central in order to compensate for the data gaps resulting from access difficulties. Finally, the research also benefitted from the observation of five multi-day events convened for the purposes of monitoring and coordinating the new agenda, and of negotiating appropriate measurement strategies. Such observations played a crucial role to make sense of and interpret the data generated through the interviews, and to check the validity of emerging insights into the organizational structures and the social relations underpinning goal-setting and goal-monitoring efforts. Additionally, the observation of such events proved important in order to identify (and get in contact with) potential interviewees, and to build a relationship of trust with them.

In terms of analysis, this dissertation relies on a *case-centric variant of process tracing* (Beach and Pedersen, 2013) and is oriented towards the specification of the rationales, step sequence and enabling conditions that explain how a certain actor affects and shapes a given outcome. This is line with the mechanism-based approach outlined by Checkel (2015) for the study and theorization of the role of IOs. This approach proves particularly appropriate to understand how inter-organization and intra-organization negotiations (informed in turn by distinct institutional cultures) affected the final wording and content of the SDG4. Importantly, this approach goes beyond the description of the process by providing the foundations to extrapolate the results beyond the bounds of the SDG4 case, and thus contributes to a more fine-grained understanding of transnational norm-building and agenda-setting.

## 1.5. Structure of the dissertation

This dissertation is organized as follows. The next chapter (*Chapter 2*) offers an historical overview of the use of international commitments and aspirational targets in the education field – in order to contextualize the object of study of this research. The chapter discusses early commitments to the universalization of primary education, the EFA experience and the MDGs. The theoretical framework informing the research is presented in *Chapter 3*, which brings together three distinct areas of scholarship, namely, international norm-building processes, the construction and impact of global indicators, and the relational character of global governance – with the objective of articulating the theoretical interface between global goals and international organizations. *Chapter 4* presents the empirical strategy on which this dissertation relies. The chapter discusses both the data-gathering strategy used to generate the corpus of data informing the dissertation, and the analytical strategy employed to make sense of it.

I then move on to the presentation of the main results of the research. *Chapter 5* discusses the findings relative to Research Objective 1, and reconstructs and examines the process through which SDG4 was negotiated. It does so from two distinct entry points, which correspond to two separate sources of debate identified over the course of the research – namely, the very architecture of the debate, and the intersect between the post-2015 process and the so-called learning turn. *Chapter 6*, in turn, grapples with Research Objective 2. The chapter focuses thus on the organizational dynamics behind the production of metrics corresponding to Target 4.1. – including the process through which specific data producers were selected, a reporting protocol was designed, and consensus on a data-harmonization strategy was found. The chapter places particular emphasis on the repercussion of such developments over the status and relationship between the UIS and a range of assessment producers.

Finally, the last chapter of the dissertation (*Chapter 7*) outlines the main conclusions of the study. To this purpose, the chapter recapitulates the main findings of the research, and discusses the potential contributions of the dissertation to two separate research agendas and literatures – namely, the practice of norm-setting and the production of global data. In the chapter I also offer some reflections on the methodological and analytical challenges entailed in the study of consensus-making in polycentric deliberation arrangements. The chapter concludes with an overview of potential venues for further research.

# Chapter 2

## Institutionalizing aspiration: a historical perspective on education goals and targets

**Introduction**

In a 2007 essay, Phillip W. Jones observed critically that "every 10 years or so since the end of World War II, the international community engages in a ritualized drama around the twin themes of universal education and the elimination of illiteracy" (2007, p. 522). Sometimes derided as a mere instance of unrealistic, "aspirational planning" (Lewin, 2007), the reliance on normative appeals to transformation remains indeed a fixture of the global education field. Time and again, international organizations, donors and governments engage in displays of collective commitment towards the full realization of the right to education. Such statements of shared aspiration are generally announced in high-profile events heralded as historical encounters, and appear to adopt a variety of forms – including international declarations, pledges and global goals and targets. The reliance on aspirational commitment thus exhibits both an enduring quality and a mutating nature.

This chapter offers a retrospective account of the use of international commitments (i.e., statements of shared purpose) in the education field. It focuses primarily on goal- and target-setting as two of the most recurrent expressions of collective commitment, although attention is also paid to the conferences and high-profile gatherings where high ideals are routinely invoked. It should be noted that, rather than assessing the ultimate impact of global goals, declarations or targets, the chapter intends to bring into focus the organizational dynamics driving the rise and consolidation of such practices. In so doing, the chapter aims to identify the repertoires of action, codified procedures and institutional forces and tensions that have shaped the negotiation of Sustainable Development Goals.

The chapter is organized as follows. The next section maps the early history of goal-setting, with a focus on the pledges to universalize primary education. The chapter then turns to examine the origins and evolution of the Education for All (EFA) movement as one of the most ambitious attempts at securing a global consensus around a common set of principles and priorities. Particular emphasis is put on the two major conferences through which EFA consolidated as a global agenda, that is, the World Conference on Education for All (Jomtien, 1990) and the World Education Forum (Dakar, 2000). The chapter then addresses some of the most relevant transformations in the global architecture of education brought about by EFA – most notably, the emergence of new actors or the transformation of their mandates. Finally, the chapter gives an overview of the origins and negotiation of the Millennium Development Goals (MDGs) and their consolidation as a guide for international development efforts in the area of education.

## 2.1. Origins of goal-setting in education

This section reviews some of the earliest international declarations and instances of global goal- and target-setting in the education realm. While some of these episodes date back to the first half of the 20th century, many of the principles advanced by such agreements continue to resonate today. More importantly for the purposes of this dissertation, these early exercises of collective commitment set in motion a process of institutionalization of goal-setting practices that proved determinant in the years that followed, shaping subsequent conferences and goal sets. This section investigates the rationales and contextual factors that motivated this first cycle of pledge-like commitments and targets, as well as the progressive routinization of these formats.

### 2.1.1. Historical antecedents

While a number of scholars coincide in observing that goal- and target-setting have a long and well-established tradition in the education realm, systematic accounts of the early days of such practices are relatively scarce. A notable exception to this is the work of Kenneth King, in particular where he digests what he terms "the pre-history" of global education targets (King, 2016). Here, the author offers a unique retrospect of the period from the 1920s to the 1960s, i.e., the period that preceded the so-called golden age of UN conferences and targets, for which there is more profuse literature (cf. Emmerji, Jolly and Weiss, 2001, *i.a*.).

As documented by King (2016), a turning point in the history of the global coordination of education aspirations was the establishment of the International Bureau of Education (IBE) in 1925. With a membership that included governments and international organizations, the IBE became one of the first organizations to promote an international perspective on education, with a particular emphasis on curriculum. More importantly for the purposes of this dissertation, as a result of the Conference on Public Education held in Geneva, in 1934 the IBE issued a pledge-like *Recommendation* calling for the universalization of basic education. As highlighted by King (2016), the Recommendation bears a striking similarity to SDG4 and can be seen as the first instance of goal-setting in education (see also Birdsall and Vaishnav, 2005; and Clemens, 2004, for a similar argument on the inaugural nature of the 1934 declaration)[4]. The statement was indeed the

---

[4] While this represents the first prescriptive stance produced as a result of an international conference, education conferences themselves cannot be considered as an entirely new phenomenon. As noted by McNeely (1995), the institutionalization of most European education systems was greatly influenced by a number of international conferences celebrated during the second half of the 19th century – including those celebrated in the context of the international exhibitions, but also education-specific events such as the 1889 Education Congress and Exhibition in Paris. Similarly, Chabbott (1998) lists a number of major conferences on educational development taking place during the late 19th century and the early 20th century – including the 1876 International Congress on Education in Philadelphia, the 1880 International

first of what would eventually become a series of 65 Recommendations, published between 1934 to 1999. In these pronouncements, the IBE advocated for a wide range of transformations in education, dealing with issues such as the principle of compulsory education, the professional training of teachers, a wide variety of curriculum-related themes, the financing of education, infrastructure and equipment, among other measures (IBE, 2020). In this sense, the IBE recommendations prefigured the expression of aspiration as one of the working modalities of choice among the global education community.

Also according to King's (2016) account, other notable precursors of goal-setting practices were the inclusion of an "education-related" human right in the 1948 Declaration, and importantly, the publication of the first World Survey of Education in 1955 and 1958. The latter was in fact the first text explicitly repositioning the education-related human right as an "educational goal common to all peoples" and framing Universal Primary Education (UPE) as a "world problem". As documented by Clemens (2004), declaratory activity continued during the 1950s with a series of UNESCO Regional Conferences on Free and Compulsory Education – celebrated in Bombay, Cairo and Lima. To be sure, these documents always refrained from establishing specific targets (let alone *quantifiable* targets), and from setting up timelines and deadlines for the consecution of these (broadly formulated) goals and ambitions. However, they were instrumental in cementing a certain credence in the transformative power of aspirational declarations.

### 2.1.2. Early experiences with target-setting: pledging to UPE

It was during the 1960s that goal-setting became definitively entrenched as a distinctive feature in the education-for-development realm – this time, together with target-setting practices. In a six-year span, up to four sets of targets were adopted as a result of a series of regional education conferences of ministers of education, organized under the auspices of UNESCO. These efforts were examined by Fredriksen (1981) in an early attempt to take stock of the progress made as a result of such collective commitments. Thus, in 1960, and as a result of the Karachi conference, 1980 was established as the target year for the achievement of UPE in a number of Asian countries. This *goal-cum-target-year* structure in fact became a template for the conferences that followed. The same objectives were established for a number of African countries as a result of the 1961 conference in Addis Ababa. A similar agreement was reached in the 1966 Tripoli conference, in this case with a focus on Arab states. Latin American countries also agreed on a similar target in the 1962 Santiago conference, although on this occasion, the deadline was established for

---

Congress on Primary Education in Brussels, and the 1908, 1912 and 1922 International Moral Education Conferences – held in London, the Hague and Geneva, respectively.

1970 instead of 1980[5]. The collective commitment to achieve UPE by 1980 was proclaimed yet again in 1970, when the UN General Assembly launched the International Development Strategy for the Second UN Development Decade. The strategy included the "enrollment of all children of primary school age" as one of the goals to be achieved by the end of the decade – thus constituting the very first instance of a global goal tied to a target date (Clemens, 2004).

The political and organizational dynamics behind the planning and management of these conferences, as well as the occurrences that shaped their final outcomes, remain something of a blind spot in the academic literature. It is thus difficult to ascertain whether the blueprint-like quality of these sets of targets is the result of a deliberate and coordinated effort on the part of the conference conveners, or rather the product of a less-strategic behavior involving subtler isomorphic forces. However, some authors have conjectured on the contextual factors and conditions that made possible such events and that might explain their high levels of visibility. As suggested by McGrath's (2018) overview of the history of education and development, UNESCO conferences and the first sets of international education targets must be understood in relation to the specific line of thinking that characterized the first stage of development[6]. This first period, running from the 1950s to the 1960s, was very much influenced by Keynesian ideas and thus characterized by a great confidence in the State as a key actor, if not the main actor in development. Also, education (understood primarily as *formal schooling*) was increasingly perceived as a necessary condition for economic progress, with the economics of education taking its first steps as a discipline and making a compelling case for education as a "productive investment". Against such a context, the reunion of a large number of state representatives (specifically, minsters of education), the emphasis on primary education, and the development of educational plans associated with education targets, were essentially in consonance with the spirit of the time. Other factors that contribute to explain the success and prominence of regional conferences were the (then recent) political independence of a number of countries willing to take part in exercises of international cooperation, and the establishment of a range of bilateral development agencies and international volunteer schemes eager to support the educational planning needs associated with the new targets (King, 2016; King and McGrath, 2012). Jones (2007) has observed that the celebration of such conferences needs to be understood in relation to a sense of optimism that permeated through the development community – noting that it is not by chance that these first education regional conferences coincided

---

[5] Note that the level of ambition contained in these sets of goals is more limited than some accounts have suggested. Thus, the agreement reached in Tripoli called for universal enrolment – rather than completion. It is also worth noting that, in some cases, the level of ambition shrank with the passage of time. The target set in Karachi, for instance, was watered-down in the Conference of Asian Ministers of Education celebrated one year later in Tokyo – with the region being divided into three groups, some of them with less ambitious targets. Similarly, while the Santiago conference initially established 1970 as a target line, this was pushed backed to 1975 in the Bogota conference held in 1963 (McGrath, 2018).

[6] Understood as the international endeavor to promote social and economic progress.

with the first UN Development Decades, "when hopes were high that the problems of underdevelopment could be tackled within such a [*short*] timeframe" (p. 323).

Finally, there are also those who believe that the portrayal of education targets as the *sole* or main product of regional conferences, is a misleading simplification. As Bray (1986) has argued, the conference/target equation owes much to the mystification of UPE that took place only after the conferences. According to the author, the fascination with UPE among a wide range of development agencies would have led to a certain misrepresentation of regional conferences as UPE-centered conferences. However, this emphasis on UPE neglects the wide variety of topics discussed in these events. According to Bray (1986), "The universalization was only one item on the agenda of the meetings, and was given less prominence in some conferences than in others […] In subsequent years, reporting has emphasized the UPE resolutions far more than the others" (p. 149). In any case, such dynamics are revelatory in their own right – for they are ultimately indicative of the power of succinctly-put goals and their association with quantifiable objectives, particularly when framed as moral imperatives.

*2.1.3. The routinization of aspirational commitment*

Goal- and target-setting practices continued during the 1970s and 1980s, even though this period witnessed a deceleration of this specific modality of normative activity – see Table 2.1 below for a synthesis. This "second generation" of goals and declarations of principles included the agreements aiming at UPE, but also at gender parity in education (Clemens, 2004). It is important to note that, whereas the first round of education targets generally aimed to achieve UPE by 1980, the target year was moved to 2000 for most of the education targets adopted during this second wave of goal-setting. While inevitable, such date-shifting is unlikely to have contributed to the credibility of goal-setting as a realistic, practical exercise.

**Table 2.1. Education goals and international commitments, 1960-1989**

| Date of approval | Forum | Goal or collective commitment | Target date |
|---|---|---|---|
| 1960 | UNESCO Meeting of Representatives of Asian Member States on Primary and Compulsory Education, Karachi | UPE | 1980 |
| 1961 | UNESCO Conference of African States on the Development of Education in Africa, Addis Ababa | UPE | 1980 |
| 1962 | UNESCO Conference of Ministers of Education and those Responsible for Economic Planning, Santiago de Chile | UPE | 1980 |
| 1966 | UNESCO Conference of Ministers of Education and Ministers Responsible for | UPE | 1980 |

| | | | |
|---|---|---|---|
| | Economic Planning in the Arab States, Tripoli | | |
| 1970 | International Development Strategy for the Second UN Development Decade, New York | UPE | 1980 |
| 1975 | World Plan of Action for the Implementation of the Objectives of the International Women Year, Mexico City | Gender parity in education | 1980 |
| 1979 | UNESCO Conference of Ministers of Education and Those Responsible for Economic Planning of Member States in Latin America and the Caribbean, Mexico City | UPE | 2000 |
| 1980 | International Development Strategy for the Third UN Development Decade, New York | UPE | 2000 |
| 1980 | Programme of Action for the Second Half of the United Nations Decade for Women, Copenhagen | Gender parity in education | 1995 |
| 1985 | Nairobi Forward-Looking Strategies for the Advancement of Women Towards 2000 | Gender parity in education | 2000 |

Source: Adapted from Clemens (2004).

In addition, conferences on the educational development theme continued to be organized during these years – with the resuming of regional conferences of ministers of education, a number of UNESCO-sponsored literacy conferences, but also a range of donor meetings instigated by private organizations such as the Rockefeller and Ford Foundations, or the UNESCO International Institute for Educational Planning (IIEP) (Chabbott, 1998). Importantly for the purposes of this dissertation, this routinization led to a certain codification of the "basic recipe" or blueprint for such events, and instilled a sense of familiarity with the conference format among a number of regular attendees.

Thus, and even if the impact of these target-setting exercises remains open to discussion, the significance of such events lies in the fact that they contributed crucially to the routinization and, consequently, the institutionalization of goal- and target-setting as a development strategy. As noted by Jolly, Emmerji and colleagues (see Jolly, Emmerji and Weiss, 2009; and Jolly, Emmerji, Dhai and Lapeyre, 2004), the entrenchment and normalization of educational goal- and target-setting inaugurated a tradition that would eventually become one of the UN procedural hallmarks, and one of its better-known innovations (see also Roberts, 2005; and Vandemoortele, 2008, for a comment on the early instances of target-setting through the establishment of Development Decades). During the 1960-2000 period, more than 50 quantifiable and time-bounded goals were adopted within the UN system – frequently, amidst a general sense of skepticism among other development quarters as a consequence of what appears to be an uneven record of success (Jolly et al., 2009).

It should also be noted that the enthusiasm for the *conference* format suggested by the list above did not occur in a vacuum. The development was coherent by the institutionalization of the conference system within the broader UN system. As reported by Emmerji, Jolly and Weiss (2001), during the 1970s the UN launched a series of thematic conferences with the aim of raising awareness of global challenges and galvanizing action around them. The authors note that such events experimented a rapid process of institutionalization, thus consolidating a basic template that included the establishment of a temporary secretariat, a series of preparatory committees and a range of standard outcomes (most notably, final declarations and programs of action). As argued by Haas (2002), global UN conferences became one of the most successful institutional innovations advanced by the UN. Despite the mounting skepticism regarding their ultimate effectiveness in bringing about tangible change and yielding durable effects (see Fomerand, 1996, for a discussion), they remained a highly popular operational modality within UN quarters. As a product of the intensification and institutionalization of such gatherings, sectoral conferences consolidated as a distinctive UN practice to the point of acquiring a ritual quality.

## 2.2. Aiming for Education for All

The advent of the Education for All (EFA) agenda in 1990 has been widely recognized as a major breakthrough – for it consolidated the centrality of education themes within development quarters, and signaled the emergence of a new consensus among the major international organizations and donors operating in the education field. This section traces the origins of the initiative as well as its evolution through the two major conferences, i.e., the World Conference on Education for All (WCEFA – Jomtien, 1990) and the World Education Forum (WEF – Dakar, 2000) through which EFA was consolidated as a global agenda.

### 2.2.1. A fragmented field

In order to understand the transcendence of the EFA agenda, it is important to bear in mind the radical transformation experimented by the global education field during the late 1970s and 1980s. This period witnessed the rise of a range of multilateral agencies which effectively disputed the leadership and centrality of UNESCO as the leading UN agency in education – a process that has been extensively documented by Mundy and Jones in a number of works. Thus, during the 1970s, the centrality of UNESCO in the provision of technical and operational assistance declined significantly. At least when it comes to research and education programming and planning, UNESCO was

progressively overshadowed by the World Bank[7] (Jones with Coleman, 2005; Mundy, 1999).

Such a shift was largely driven by the climate of political turbulence that surrounded UNESCO during the 1970s – and which eventually translated into a significant downsizing of its budgetary basis. Unlike other multilateral organizations, and partially as a consequence of its specific voting and decision-making structure, UNESCO was unable to remove or insulate itself from the demands of an increasingly cohesive Third World bloc calling for a reformulation of the North-South relations and advocating a reorganization of the UN. In an attempt to respond to developing countries or to appease those countries, UNESCO introduced a number of philosophically oriented themes (i.e., endogenous development and the cultural dimensions of learning) while favoring the development of regional education agendas (Mundy, 1998, p. 39). The development of an increasingly expansive and diffuse program, coupled with what was came to be perceived as an excessively bureaucratic and slow administration, translated into a generalized "loss of confidence" in the institution (Mundy, 1999). In addition, and perhaps more importantly, it offered other organizations the opportunity to carve a role for themselves in the education realm.

Thus, the World Bank, which had long relied on UNESCO's technical advice and support, started to develop an expertise of its own – a move that added to its financial dominance (Jones, 1992). Similarly, UNICEF put an end to a tradition of collaboration with UNESCO in order to privilege a more targeted, hands-on approach. Through the development of the so-called "basic needs approach", these two organizations succeeded in striking a balance between the calls for change put forward by developing countries, and the pragmatic approach to development supported by Western governments (Mundy, 1999). The growing prominence and authority enjoyed by these two organizations had a direct impact on UNESCO's reputation, with up to three countries (the US, the UK and Singapore) leaving the organization in a two-year period (1983-1984) over criticism of UNESCO's politicization and mismanagement[8]. This withdrawal entailed a 30% reduction of UNESCO's (already strained) budget. It is in this context of scarce resources that normative and persuasive activity came to be associated with UNESCO as its specific modus operandi. A division of labor between "action" and "reflection" emerged gradually – with the UNESCO program privileging normative suasion, and the rest of UN agencies becoming the main providers of technical and financial assistance (Jones with Coleman, 2005).

---

[7] While the World Bank had entered the education scene in the 1960s, education remained initially a peripheral theme in the organizational agenda and its programmatic priorities.

[8] To be sure, the withdrawal of such countries was largely driven by reasons external to the education field – it was catalyzed greatly by UNESCO's support to the "New World Information and Communications Order" mischaracterized as a threat to the freedom of press (Mundy, 1999). Similarly, and as noted by Jones, part of the mounting criticism faced by UNESCO during this period was essentially "an early expression of the anti-UN stance of neoliberalism" (p. 528).

*2.2.2. The origins of the World Conference on Education for All*

Even if by the late 1980s normative activity had been consolidated as UNESCO's forte, the WCEFA idea was not born in the quarters of this organization. At least in its embryonic form, the idea of a world education conference appears to have been first envisaged in 1987 by UNICEF's Executive Director, James Grant, and conceived as an attempt to replicate and reinforce the success of different health-related initiatives launched by this organization in the late 1970s. The process through which such an idea materialized into the WCEFA celebrated in Jomtien in 1990 has been extensively documented by Chabbott (1998, 2003) and Jones (2007). Despite emphasizing different aspects of the process, both authors concur in noting that the different organizations that eventually joined the initiatives did so for different reasons and with different (and sometimes irreconcilable) expectations. As discussed below, the generalist tone of the agreement reached in the conference owes much to this disparity of rationales – and so does the uneven record of achievement of the EFA agenda.

The leading role exhibited by UNICEF in the conception of the WCEFA idea needs to be understood in relation to the Health for All movement launched in the 1987 Alma-Ata conference, convened by UNICEF and the WHO. The Health for All agenda laid the foundation of two highly successful initiatives that would greatly enhance UNICEF's reputation in the following years – a campaign disseminating basic information on child mortality, and a child immunization campaign. As per Jones's (2007) account, the success of these two simple and relatively inexpensive initiatives led Grant to believe that, in the context of a sufficiently-resourced campaign, UPE was an equally achievable goal. In addition, within UNICEF quarters, education came to be perceived as a necessary means to realize the Health for All agenda. It is against this background that Grant embarked on the quest to convince other UN agencies of the potential of a high-level event – with what Jones describes as "an almost missionary zeal" (2007, p. 527).

UNESCO was an early enthusiast of the WCEFA idea for reasons relative to the loss of prestige of the organization – which by the mid-1980s had reached an all-time low. According to Chabbott (2003), Federico Mayor (the Director-General of UNESCO) was quick in appreciating the potential of a high-profile, apolitical event as a privileged opportunity to regain some of the organization's lost centrality. A world conference was also perceived by UNESCO as an opportunity to improve its reputation by association with the two other major players in education (Chabbott, 1998).

Conversely, the World Bank was more of a late, hesitant adopter. Still according to Chabbott (2003), the World Bank only joined in 1988 as a result of Grant's mediating role. Thus, despite the organization's initial reservations, Grant convinced the World Bank President, Barber Conable, of the need for a conference as a means to develop the intellectual consensus necessary for realizing the universalization of basic education. Additional detail on the participation of the World Bank is provided by Heyneman (2009) in his inside story of the internal debate generated by the proposal. According to

Heneyman's account, the project was initially opposed by the Vice President for Central Projects (Warren Baum) on the grounds that the World Bank could not commit itself to prioritize a specific sector. Jones (2007), in turn, draws attention to the emergence of internal concerns over the feasibility of the agenda, noting that:

> Senior management were fearful that the Jomtien initiative had considerable potential to commit the international community to targets that the Bank itself would or could not commit to […] Even its education chiefs wanted no part of an EFA plan that would force countries into unachievable targets or undeliverable claims on both domestic and international finance. (p. 533)

However, the Director for the Education Department (Aklilu Habe) became convinced of the potential of an interagency project, and the idea of a major conference came to be increasingly perceived as an opportunity for the Bank to persuade the international community of the need prioritize primary education. By the late 1990s, this was a particularly pressing need given that the World Bank policy in education had essentially experienced a U-turn – shifting from an emphasis on secondary, technical and higher education, to the elevation of primary education as an overriding priority (Jones, 2006)[9]. The WCEFA was thus repurposed as a means to vindicate basic education as a development priority for both donors and borrowing countries. After consultation with the education sector chiefs, an agreement was reached to support the project by providing financial and staff resources, while avoiding any kind of commitment in terms of lending priorities. The World Bank sponsorship of the WCEFA project was thus made conditional on the avoidance of financial targets – expressed either as a percentage of GDP or of the Bank's lending in education (Heyneman, 2009; Jones, 2007).

The UNDP was a late-comer to the project, becoming a sponsor in late 1988. As noted by Jones (2007), it remained a "reluctant partner" somehow forced to join the initiative in its function as the lead UN agency for development, in view of the (potential) implications of the conference for UN funding for basic education. Thus, the organization kept a low profile during the conference, frequently adhering to the World Bank proposals, and its intellectual or strategic contribution to the initiative appears to have been negligible.

In preparation for the conference, the four agencies established the Inter-Agency Commission (IAC), led in practice by a small Executive Secretariat located at UNICEF quarters in New York and headed by Wadi Haddad from the World Bank – with the assistance of one official from the other three agencies. The IAC was also advised by two additional bodies – an Executive Committee and an International Steering Committee.

---

[9] This turnaround was largely the product of the growing ascendancy enjoyed by the proponents of the rates of return within the World Bank's education sector. Eventually, such measures became the dominant criteria in the Bank's education planning efforts, contributing to the rise of what Heyneman (2003) called the "short policy menu" – which required low-income countries to shift public expenditure to primary education, while financing higher education on the basis of user fees.

While the specifics of these bodies are beyond the scope of this chapter, the relevance of IAC's tripartite structure lies in the fact that it came to decisively influence the EFA architecture set up at Jomtien – especially the division of labor between an International Consultative Forum, its Secretariat and a Steering Committee (Little and Miller, 2000). It is also worth noting that, prior to Jomtien, the IAC Executive Secretariat engaged with a number of consultative hearings – including meetings with governments, non-governmental organizations (NGOs) and a series of nine regional consultations (Little and Miller, 2000). The reliance on constituency-specific and regional consultations would also become a fixture of the EFA machinery.

### 2.2.3. An improbable agreement?

The preparation of the WCEFA succeeded in building momentum and attracting the interest of a wide range of constituencies. Participants at the conference included around 150 national delegations, representatives from more than 30 IOs (including UN agencies and other inter-governmental organizations operating at a regional level, as well as regional banks), and around 125 delegates from NGOs. In addition, while the four convening agencies remained the main sponsors[10], the conference secured additional support of a number of co-sponsors and associate sponsors[11] (including a number of OECD governments, the US and Canadian bilateral agencies and a number of UN agencies) (IAC-WCEFA, 1990).

Despite this wide participation, the planning and execution of the conference (as well as the drafting of the ensuing Declaration), was very much the product of a few organizations (most notably, the conference conveners) – whereas the role of governments (especially, those from developing countries), researchers or NGOs was purely anecdotal (King, 2007). It is quite revealing in this regard that, even if these constituencies were present in the meetings prior to the conference (and national delegations included both government and civil society representatives), the drafting committee was drawn exclusively from multilateral agencies. As documented by King (2007), the draft discussed in Jomtien evolved directly from an initial text prepared in the context of a UNICEF consultation in February 1988, with the conference providing only an opportunity for its *revision*.

However, the limited number of organizations directly involved in the drafting should not imply that the final agreement was reached without friction. As noted in most of the

---

[10] In accordance with the nomenclature used by both official documentation and academic literature, the terms EFA sponsor and EFA convener are used interchangeably in this dissertation. This refers to the main sponsors of EFA conferences, that is, UNESCO, UNICEF, the World Bank, UNDP and, since 1998, UNFPA.

[11] Co-sponsors contributed with half the amount brought by sponsors, whereas associate sponsors contributed around 20% of this.

available accounts, different participants emphasized different themes, or pushed for certain frameworks perceived as less acceptable by other attendees. As documented by Chabbott (2003), King (2007) and Torres (2000), the main bone of contention was the relative priority given to primary education. Whilst the World Bank privileged formal primary education, UNICEF was more open to non-formal education and emphasized the need to pay greater attention to early childhood development and initial education, and UNESCO placed greater priority on the education of adults and out-of-school youngsters. Despite the World Bank's insistence on the UPE agenda, the prioritization of primary education was rendered impossible by the role of a wide range of professional constituencies, each one of them concerned with a specific education sub-sector and keen and committed to make sure that their priorities were captured in the drafts prepared before and during the conference. As discussed below, this is in fact one of the main reasons behind the use of the all-encompassing notion of *basic education* in the final Declaration – and the limited number of references to specific education levels or modalities of education delivery.

In her extensive analysis of the WCEFA episode, Chabbott argues that one of the main factors that made it possible to reach a consensus despite the disparity of the agendas was the existence of a "basic recipe" guiding the planning and execution of the conference (1998, 2003)[12]. According to the author, the IAC relied heavily on pre-existing models (as those having oriented prior conferences) in order to ensure the WCEFA's conformity and adherence to internationally-accepted standards of fairness and objectivity. This would have prevented major clashes – at least over procedural matters. When it comes to the very content of the agreement reached, the author has noted that the consensus made at the WCEFA tapped into the UN human rights instruments already enjoying a strong buy-in on the part of most member states. Overall, the goal-setting and prescriptive activity having taken place over the previous decades, along with the familiarity of many of the WCEFA attendees with both the conference and the international declaration formats, played a key role in facilitating the success of the conference:

> Many of the delegates to the regional consultations that produced the Declaration of Education for All in just six months were, therefore, no strangers to each other; they had met and discussed with each other many of the same issues, dozens of times in their careers. In addition, many of the principles underlying the Declaration came directly from documents with which the delegates were already familiar: the Declaration of Universal Human Rights, the International Covenant on Economic, Social, and Cultural Rights (Article 13, 1966); the Declaration (1959) and Convention on the Rights of the Child (Articles 28 and 29, 1989); and a half-dozen other conventions and agreements. Thus by

---

[12] Certainly, the emphasis of Chabbott's accounts (1998, 2003) on consensual dynamics is not always consistent with other authors' emphasis on the inter-organizational infighting that took place in Jomtien. This cannot be disconnected from the fact that Chabbott's description of the WCEFA episode is explicitly put at the service of World Culture Theory. The reliance of the WCEFA conveners on a series of heavily-scripted frameworks and blueprints is thus presented as evidence of the constraining effect of the repertoire of action put forward by the so-called World Culture (a highly rationalized, Westernized and scientized worldview putatively playing a dominant role, according to the proponents of such theory – cf. Thomas et al., 1987).

imitating the format of earlier conferences and invoking ideals that were taken-for-granted in international discourse, the organizers mobilized both mimetic and normative pressures to ensure their success. (p. 213)

As noted by Chabbott (1998), this is all the more relevant considering the low level of intellectual and professional consensus that characterized the educational development field by the late 1980s – particularly as opposed to "homologue" sectors such as health, or as opposed to the early days of education conferencing. In a field then fraught with a number of deep-running, ideologically-driven tensions, the prospect of reaching a consensus represented a particularly challenging endeavor and, consequently, was particularly dependent on the existence of broadly-accepted procedural scripts.

### 2.2.4. The end products of the WCEFA

The main products of the WCEFA included a normatively-oriented World Declaration on Education for All along with a Framework for Action (FFA) (IAC-WCEFA, 1990). When it comes to the former, different authors have drawn attention to its generalist overtone – which contrasts greatly with the emphasis placed on UPE by prior goal-setting exercises. The relative vagueness of the final wording has been extensively discussed by Torres (2000), who observes that the notion of basic education in fact encompassed different (and sometimes conflicting) views and priorities of those involved in the drafting. A similar point has been raised by King (2007), who notes that:

> The Conference was much less about the conventional categories of primary, secondary, technical and vocational, and higher education than about the new would-be inclusive concept of basic education or the rather vague notion of "basic learning needs". Indeed, surprisingly, with the exception of primary education, none of these traditional categories is named at all in the entire document, with the exception of higher education which is mentioned just once (WCEFA, 1990). (2007, p. 379)

The relevance of such dynamics lies in the fact that, as developed later, the reliance on blurry concepts would eventually become a patterned modus operandi in global goal-setting exercises. It is also important to bear in mind that the reliance on *basic education,* a term subject to multiple understandings translated eventually in the inconsistent use of the term on the part of the EFA partners.

The vagueness of the Declaration was in addition compounded by the striking absence of targets – that is, time-bound, quantifiable objectives. This was much to the dismay of UNICEF's Executive Director, James Grant, who saw them as essential if the Declaration was to secure the necessary support from donors. However, the drafting committee judged it more appropriate to encourage countries to set their own targets, suggesting only a list of dimensions (early childhood care, adult literacy, UPE, and so on)[13] that countries could

---

[13] While these dimensions have sometimes been mistakenly portrayed as goals, they were in fact less prescriptive than such labeling might suggest. The FFA expressly noted that "Countries may

take into account when devising national strategies (Chabbott, 2003). Eventually, and as a result of Grant's pressure, two of the suggested targets were explicitly set for the year 2000, thus acquiring a target-like quality. These targets were universal access and completion of primary education, and a reduction of the adult illiteracy rate to one-half of its 1990 level[14]. Such development is relevant in that it prefigured one of the main sources of tension that would re-emerge in subsequent goal-setting exercises – that is, the trade-off between establishing universal goals able to signal priority issues, and allowing countries to establish their own objectives appropriate to their reality and institutional, economic or political capacity.

If the Declaration was essentially an expression of high (and somewhat vague) ideals, the accompanying FFA constituted an action-oriented piece, in which the means and strategies to realize the EFA promise were specified. Participating countries therefore committed to take a series of steps in order to ensure progress toward the realization of EFA goals – most notably, the delineation of national strategies and plans for action. In addition, the FFA devised a mechanism to operate at an international level oriented at serving follow-up action and at fostering a "spirit of co-operation" among countries, multilateral and bilateral agencies, and NGOs. As noted by Little and Miller (2000), building a consensus around this area proved a particularly challenging endeavor, and rendered evident a number of political tensions. The role of the World Bank was a first point of contention as, while some perceived that the World Bank could take the lead in the implementation of EFA (coordinating national plans and providing resources if necessary), others, including the Nordic bilateral agencies, the EU and UNESCO, advocated for a less top-down strategy. Secondly, the role of UNESCO proved contentious, as the idea of conferring this agency with a leading role was flatly rejected by the US. Through a number of side-meetings, a final compromise was reached according to which "the follow-up secretariat would be sited at and serviced by UNESCO, but it was not [*to be*] part of UNESCO" (Little and Miller, 2000, p. 7).

The follow-up mandate was thus given to an International Consultative Forum on Education for All (known as the EFA Forum) expressly created for such purposes. The Forum was integrated by representatives of the five sponsoring agencies as well as of bilateral aid agencies, governments and NGOs, and researchers – however, in practice, the bulk of its activity was led by a UNESCO-based Secretariat (Torres, 2001). Thus, whereas the Forum had only a consultative and advisory role (as it was conceived as a space for debate and exchange), it was the Forum Secretariat that was tasked with the

wish to set their own targets for the 1990s in terms of the following proposed dimensions" (IAC-WCEFA, 1990, p. 52).

[14] It is also worth noting that, even if 2000 was not expressly defined as a deadline or target year for EFA, it was increasingly perceived as such by a wide range of organizations. This was partially the result of the fact that the leaders of all four sponsoring agencies referred to 2000 in their statements during the WCEFA. As noted by Torres (2000), EFA was adopted in practice as a decade-long program – a spirit clearly reflected in the end-of-decade assessment, which very much resembled a "final" assessment.

organization of biennial meetings, providing information on EFA, and engaging in advocacy activities. In addition, a supplementary advisory group was set up in 1991 in order to complement the work of the EFA Secretariat. This was the Forum Steering Committee, which included the original EFA sponsors, UNFPA (which could eventually become the fifth convener in 1998), a number of bilateral agencies and a representative from the NGO sector. The division of labor between the Secretariat and the Steering Committee, however, remained unclear and was characterized by a number of organizational overlaps and issues of duplication. As argued by Little and Miller (2000) in their evaluation of the EFA Forum, such dynamics greatly limited the efficiency of the follow-up mechanisms:

> The Forum, the Forum Secretariat and the Forum Steering Committee were created in response to concerns raised with respect to mechanisms that were needed to follow-up on the World Conference. Jomtien was not to be an event but the initiation of an ongoing process. From the above analysis it is clear that the operations and interrelations of these mechanisms were not anticipated in fine detail. Terms of reference for each, and in relation to each other, were not worked through. All three began life as loosely organized coalitions of interests of different EFA stakeholders, with the interests of conveners predominant. Mandates were unclear. Tensions and contradictions were inevitable. (p. 12)

As noted by the same authors, the influence exerted by the Forum, and especially the Forum Steering Committee, was also limited by a lack of clarity regarding issues of organizational representativeness. Thus, it always remained unclear whether individuals attending the Forum meetings were joining in a personal capacity, or whether they were representing their organizations. Interestingly enough, while representative of the original EFA sponsors tended to see themselves as representing their institutions, those individuals nominated by bilateral agencies were far less likely to do so. It should also be noted that the progressive expansion of the Committee[15] and the incorporation of new members made it increasingly difficult for this advisory group to provide clear guidance.

*2.2.5. Assessing the impact of EFA*

In stark contrast with the high ideals invoked in the WCEFA, the progress made by the EFA agenda over the years that followed Jomtien was extremely limited. The lack of progress towards the realization of the EFA promise was in fact routinely reported during the follow-up meetings that took place during the 1990s – in particular, the periodic meetings of the EFA Forum. Prospects for the consecution of EFA became particularly bleak after the Amman Forum of 1996, where a mid-decade assessment was presented. The assessment identified a number of achievements, but highlighted that the progress was not as fast or as widespread as originally deemed possible, and that most of the

---

[15] The composition of the Steering Committee was expanded in the mid-1990s in order to incorporate a greater number of NGO representatives as well as a number of "personalities" from developing countries.

education levels or modalities other than primary education and formal schooling had been neglected and remained underdeveloped. Other assessments prepared by the end of the 1990s and oriented towards taking the pulse of the EFA movement reached similarly somber conclusions (Torres, 2000).

According to more recent accounts of the EFA project by Mundy (2006, 2010), the limited traction gained by the EFA agenda during this first period needs to be understood in relation to the fall of bilateral aid for education, but also in relation to the fragmentation of the education and development field – with different IOs guided by different interests and credos, and varying levels of technical expertise. These organizations were thus keen to instrumentalize EFA to enhance their own reputation. During the decade from 1990 to 2000, such dynamics were magnified and greatly compounded by the absence of an effective coordination mechanism, the precariousness of the following-up mechanisms (Jones, 2007), and the marginalization of the NGO sector – which counted itself among the most vocal advocates of the development of global mechanisms of accountability (cf. Mundy and Murphy, 2001)[16].

Thus, even if the differences among the WCEFA conveners had not precluded them from reaching a consensus regarding the wording of the EFA Declaration and the FFA, the uneasiness of such an alliance became increasingly apparent during the years that followed (Jones, 2007). Indeed, the EFA agenda exerted a very limited influence on the programmatic priorities and agendas of the partaking IOs, and despite the repeated calls for strengthened financial support and inter-agency cooperation, most organizations stuck to their original funding priorities. This translated into a virtual lack of funding for certain areas – for, among the better-resourced agencies, priority was almost invariably given to primary education; whereas UNESCO (the most explicit proponent of a "broad" approach to EFA) yielded limited authority and power. As noted by Chabbott (2003), the absence of economists among UNESCO ranks, as well as the limited funds available for scientific research, prevented this organization from making a compelling case for adult education and vocational training. UNICEF, in turn, selectively turned its attention to a subset of UPE – specifically girls' education. The implementation of the EFA agenda was thus largely driven by the two largest donors to education at the time, the World Bank and the United States – both of them determined to focus efforts on access to formal primary education (Heyneman, 2009).

As noted by Torres (2000), such dynamics translated into a general narrowing of the EFA spirit. Thus, the principles advanced by the EFA declaration under the banner of *basic education,* which corresponded to an "expanded" vision of education (see Table 2.2 below), ended up shrinking during the period 1990 to 2000. Thus the implementation of the EFA agenda translated into a return to the "restricted" or conventional view of education.

---

[16] This is discussed in greater detail below – see section 2.3.5.

**Table 2.2. Restricted and expanded visions of education – Torres, 2000**

| Restricted vision (conventional) | Expanded vision (Education for All) |
|---|---|
| Directed at children | Directed at children, youth and adults |
| School-based | Based inside and outside school |
| Equates with primary education or a pre-established level of education | Is not measured by the number of years of study or certificates attained, but by what has effectively been learned |
| Responds to the teaching of specific subject | Responds to the satisfaction of basic learning needs |
| Recognizes only one type of knowledge as valid; that acquired in the school system | Recognizes the validity of all types of knowledge, including traditional knowledge |
| Is uniform for all | Is diversified (basic leaning needs are different in different groups and cultures, as are the means and modalities to meet those needs) |
| Is static ("change" takes the form of periodic school and curriculum reforms) | Is dynamic and changes with the passage of time (educational reform is permanent) |
| Supply (institution, school system and administration) predominates in the definition of content and methods | Demand (students, family, social demand) predominates in the definition of content and methods |
| Focussed on teaching | Focussed on learning |
| Responsibility of Education Ministry (education as a sector and a sectoral responsibility) | Involves all Ministries and government bodies in charge of educational actions (requires multi- sectoral policies) |
| Responsibility of the state | Responsibility of the state and the whole society, thus demanding consensus-building and co- ordination of actions |

Source: Torres (2000).

However, and despite this uneven record of success, the historical relevance of the WCEFA can hardly be overstated. To start with, the EFA Declaration was the first instance of an agreement binding together both state and non-state actors. Previous examples of global goal-setting had always adopted the form of state-based declarations or convention (Faul, 2014). In addition, it became a unifying framework in a highly fractious field. Finally, the themes and principles enshrined in these products represented a departure from the emphasis on UPE that had characterized prior expressions of collective commitment.

### 2.2.6. The lead-up to Dakar

Even if progress towards EFA stalled during its first decade of life, the reliance on global meetings and statements of principle remained a fixture of the global education community. As noted by Chabbott (2013), "EFA continued going through the motions of the world conference model throughout the 1990s" (p. 41). Thus, three global meetings were held with the objective to monitor progress towards EFA. These were the 1991 Paris

Meeting, the Summit for the Nine Most Populous Countries of 1993 (celebrated in New Delhi), and the 1996 Amman Forum, where, as noted above, the results of the mid-decade review were presented and discussed[17]. This was topped by a series of four high-level seminars targeted at ministers and policy-makers and co-organized by UNICEF and the IIEP-UNESCO.

In addition, declarative activity in relation to education continued to take place in UN conferences beyond the education realm. A particularly significant episode was the World Summit on Children, organized by UNICEF in New York in September 1990. Here the question of child education and its multiple formats and means of delivery were discussed. Other relevant events included the 1993 World Conference on Human Rights which led to the Vienna Declaration and Program of Action and which commented on the obligation of States to promote gender equality, especially in education, and the 1995 Fourth World Conference on Women, held in Beijing, which also placed high priority on girls' and women's education and training as transformative. It is also worth noting that "Universal and equitable access to education" featured as one of the 10 commitments advanced by the 1995 World Summit for Social Development of Copenhagen, and in addition even the World Food Summit of 1996 discussed strategies to support general education (Mundy, 2006; Unterhalter, 2014).

While, as some observers noted, EFA meetings failed to maintain the sense of momentum around the EFA project (Faul and Packer, 2015), their importance should not be hastily dismissed. To start with, the EFA meetings were instrumental in creating a number of "weak ties" in an otherwise sparse network populated by a number of sharply-bounded organizations with a tendency towards self-referentiality. In addition, some of these encounters were the cradle of initiatives that would eventually contribute to a certain restructuring of the field. This was specifically the case for the E-9 Initiative, born out of the New Delhi meeting. While the original aim of the E-9 Initiative was simply to foster collaboration among the nine most populous (developing) countries in the world, the project took on a life of its own with the nine countries taking control of the initiative and establishing its own Secretariat, led by ministers of education rather than by donors or IOs (Little and Miller, 2000).

The cycle of global summitry intensified during the last two years of the first EFA decade – motivated in part by the unsatisfactory nature of the mid-decade assessment presented in the Amman Conference in 1996 (Jones, 2007). Thus, in the light of sharp criticism directed towards the quality of the assessment, the EFA Forum and the Steering Committee started working on the planning of the EFA End-of-Decade Assessment as well as on the 2000 World Education Forum, to be held in Dakar (Little and Miller, 2000).

---

[17] For this reason, an effort was made to secure the presence of a larger and more representative set of countries – something that has not occurred with the Paris and New Delhi Forums. Also, the Amman meeting was preceded by a number of regional meetings and national gatherings (Little and Miller, 2000).

Interest around EFA was also reinvigorated by the adoption of a resolution of EFA on the part of the UN General Assembly (Resolution A/52/84 on Education For All). Adopted in 1997, the resolution contained a recommendation for EFA stakeholders to provide the UN Secretary-General and UNESCO's Director-General with the information necessary to report on the overall progress up to 2000. The resolution also requested the UN Secretary-General and UNESCO's Director-General to "effectively pursue EFA" in cooperation with Member States (Ito, 2012).

In the light of these developments, the executive heads of the EFA conveners met in Geneva in March 1998, and the EFA Forum established an Inter-Agency Global Technical Advisory Group integrated by officials from the five EFA sponsors. In addition, a number of national and regional reports were prepared and shared at a series of five regional meetings[18] and 10 regional consultations. This was facilitated by up to 11 Regional Technical Advisory Groups created to support and coordinate assessment efforts at a country level[19] (Ito, 2012). These constituted the basis of the EFA 2000 Assessment, a synthesis of the progress made by countries towards the achievement of EFA, along with an evaluation of the financial contributions of EFA agencies. With a global scope, the report was widely circulated before the Dakar meeting and, while it contributed to renew interest and commitment around EFA, it also confirmed that the agenda remained a challenging and elusive one (Little and Miller, 2000; Torres, 2001).

### 2.2.7. Revitalizing the EFA mandate

The 2000 World Education Forum (WEF) was thus celebrated against a much less confident or hopeful background that the one that had once characterized the pre-Jomtien run-up. Organized by the International Consultative Forum on Education for All (the EFA Forum), the purpose of the meeting was twofold; it aimed to present the results of the end-of-decade assessment, and to devise a new Framework for Action to pursue with the EFA quest (Torres, 2001).

In contrast to the WCEFA, accounts of the negotiation of the WEF are much sparser and sketchy in nature. Most available accounts focus on the "broadening" effect that the meeting had over the EFA agenda. As noted by King (2007), the meeting was seized by

---

[18] Specifically, one meeting for Sub-Saharan Africa celebrated in Johannesburg, one for Asia-Pacific celebrated in Bangkok, one for the Arab States held in Cairo, one for Europe and North America held in Warsaw, and one for the Latin American and the Caribbean region held in Santo Domingo. All of them were celebrated during 2000, with the exception of the Sub-Saharan Africa meeting, held in 1999.

[19] It should also be noted that all countries were required to establish a National Assessment Group integrated by planners, statisticians and researchers and responsible for leading assessment efforts. It is unclear how many of these groups were effectively formalized. In any case, it is worth noting that more than 180 countries participated in the assessment and 167 of them submitted a national report to UNESCO (Ito, 2012).

UNESCO as an opportunity to push for a return to the inclusive and expansive vision espoused at Jomtien but neglected over the following years. The Collective Commitment approved at Dakar did not entail significant changes in the nature of the ideals invoked at Jomtien. There were, however, a number of relevant adjustments, including the greater emphasis placed on the quality of education, the substitution of "life skills" for "skills required by youth and adults", the acknowledgement of education as a right, the statement that primary schooling should be free and compulsory, references to the need for measurable learning outcomes, greater prominence given to girls' education, and calls for the incorporation of adult education into national education planning (see King, 2007; and Torres 2001).

That said, changes in the wording cannot be mechanically equated to a change in the relative importance attached to the new formulations by the partaking organizations or the conference attendees. In fact, Torres (2001) cautions against reading too much into the exact wording of the declaration. There is a risk of over-interpreting the specifics of the text – the presence or absence of a given word might simply reflect a contingent, temporary victory of a given interest group. As noted by Torres, the text had been prepared in advance and various drafts had been circulated – to the point that "What is left open for discussion is form rather than content: replacing, deleting or adding words, moving paragraphs, or highlighting one particular idea among the whole" (2001, unnumbered).

To a great extent, Dakar operated essentially as a confirmation of the agenda set in Jomtien (in the sense that it did not advance a new approach or view regarding the purposes of education). However, Packer has noted that the significance of the WEF lies in the action-oriented nature of the agreements reached at the conference:

> In contrast [*to Jomtien*], the World Education Forum at Dakar ten years later added little to the conceptual understanding of EFA although its emphasis on the most disadvantaged in society was strong. But consistent with wider international efforts to set development targets and timelines (Social Summit in Copenhagen in 1995; the Millennium Summit, 2000) it offered a much more specific set of goals, albeit those on literacy and adult learning have not been that easy to interpret subsequently. Dakar called for an EFA global initiative and for individual countries to plan to achieve the six Dakar goals with immediate effect. It specified the need for regular international monitoring. It charged UNESCO (as had Jomtien) with playing an international coordination role. As a result of these and other recommendations, Dakar was a much more action oriented event with a body of political commitments that have helped to kick-start a range of international processes and activities over the past seven years. (2007, p. 6)

The return to the use of time-bounded targets (see Table 2.3 below) represents a case in point and warrants particular attention. The origin of this decision can be found in the technical guidelines prepared by the Global Technical Advisory Group. This group suggested a list of 18 indicators in order to support and guide the work of regional and national assessment groups (Ito, 2012; Roberts, 2005). These were largely focused on primary education and formal schooling – one third of them were directly connected to

UPE. While there is limited evidence of the reasoning that developed such emphasis, it was coherent with the growing consensus regarding the limiting effect of the lack of national and cross-national data as one of the main obstacles preventing EFA to realize its potential (Unterhalter, 2014). In any case, this move contributed to the revitalization and political traction gained by the EFA agenda over the following decade.

**Table 2.3. Dakar EFA Goals**

| | |
|---|---|
| *1* | Expanding and improving comprehensive early childhood care and education, especially for the most vulnerable and disadvantaged children. |
| *2* | Ensuring that by 2015 all children, particularly girls, children in difficult circumstances and those belonging to ethnic minorities, have access to and complete, free and compulsory primary education of good quality. |
| *3* | Ensuring that the learning needs of all young people and adults are met through equitable access to appropriate learning and life-skills programmes. |
| *4* | Achieving a 50 per cent improvement in levels of adult literacy by 2015, especially for women, and equitable access to basic and continuing education for all adults. |
| *5* | Eliminating gender disparities in primary and secondary education by 2005, and achieving gender equality in education by 2015, with a focus on ensuring girls' full and equal access to and achievement in basic education of good quality |
| *6* | Improving all aspects of the quality of education and ensuring excellence of all, so that recognized and measurable learning outcomes are achieved by all, especially in literacy, numeracy and essential life skills. |

Source: Adapted from World Education Forum (2000).

## 2.3. Dakar as a field-configuring event

While the Dakar Declaration essentially represented a reaffirmation of the principles espoused in Jomtien, the action-oriented nature of the WEF end products succeeded in catalyzing a significant restructuring of the global education field. Thus, despite the limited progress towards EFA witnessed during the post-Dakar period[20], the relevance of the WEF episode lies in the fact that it contributed to redefining the global architecture for education development. This section discusses some of these transformations – most notably, the emergence of new actors and governance spaces.

---

[20] Despite the revitalization injected by the WEF, the post Dakar-era offered a mixed record of achievement. While such debates are addressed in greater detail in Chapter 5, here it is suffice to say that after an early upsurge of official aid for education, education assistance stagnated after 2005, and started to decline after 2011. In addition, goals other than UPE and gender equality continued to receive very limited attention (Mundy and Manion, 2015).

While it is beyond the scope of this chapter to give a full account of the evolution and role of these formations, the short overviews provided in the following section are oriented at making a broader point – namely, that the practices of goal-setting and multilateral conferencing have the potential to operate as catalysts of change in transnational education governance.

## 2.3.1. A new EFA architecture

The establishment of a follow-up arrangement led by UNESCO represents one of the most relevant changes brought about by the WEF. The origins of this transformation can be found in the proposals of the so-called Futures Group (one of the Forum's working committees), which called for the establishment of a dedicated organism tasked with the follow-up of the implementation of the second "phase" of EFA. Remarkably, one of the key rationales for the creation of such a body was to avoid UNESCO becoming the de facto coordinator of the post-Dakar EFA. Thus, the original proposal advocated a coordinating body integrated by officials of all the EFA sponsoring agencies, along with representatives from the government and civil society organizations. However, the proposal was met with fierce opposition from Latin American, African and Asian representatives. Representatives from these regions were adamant that the leadership of EFA should fall on a *restructured* UNESCO – much to UNESCO's surprise. Despite the turmoil caused by such developments, the proposal was amended accordingly, and UNESCOs' Director-General (Koichiro Matsuura) committed to the requested restructuring in the final plenary (Torres, 2001).

The assignment of full responsibility to UNESCO marked a turn away from the collective leadership mechanisms put in place in Jomtien – which, as noted above, had materialized into an Inter-Agency Commission in which the EFA convening agencies shared power on equal terms. As per the Dakar FFA, UNESCO adopted a distinctly salient role, as it was tasked with the periodic convening of a High-Level Group (HLG). This was envisaged as the main coordination mechanism and expected to operate as a "lever for political commitment and technical and financial resource mobilization" (UNESCO, 2000). The group met annually for 10 years and benefited from the active involvement of a rotating group of Ministers of Education.

With the passage of time, however, the membership of the HLG expanded to include an increasingly heterogeneous group of agencies – a move that took a heavy toll on the capacity of the HLG to provide clear guidance (UNESCO, 2015). Such dynamics were compounded by the absence of clear articulation between the HLG and the Working Group – a de facto assemblage of the six goal-specific working groups established in Dakar. While the Working Group was assigned a more technical mandate (expected to complement the more political role of the HLG), in practice, the division of labor between these two bodies was never clear. As a consequence of such a lack of definition, and as

noted by Robinson (2014), the mandates of both organisms kept evolving over time – thus compromising the effectiveness of the arrangement:

> Making the two groups work as a combined mechanism to push EFA forward after 2000 proved difficult, with shifting understandings of the articulation between them. A concern to sequence the discussions coherently and ensure strong systematic messages, based on the evidence of the EFA Global Monitoring Report, led to incremental changes in timing and agendas. (unnumbered)

The lack of clarity regarding the distribution of responsibilities did not go unnoticed by UNESCO. Between 2001 and 2006, the organization issued up to three coordination plans aimed at clarifying roles – mostly to no avail[21]. Finally, in response to the limited effectiveness of the coordination mechanisms in place, a reform of the EFA architecture was initiated by UNESCO in 2011. The HLG and the Working Group were replaced by the EFA High-Level Forum (HLF), envisaged as forum of world leaders and education experts. The work of the HLG was to be complemented with the periodic convening of Global Education Meetings (GEM), attended by ministers, senior officials and technical officers (EFA-GMR/UNESCO, 2015). However, as noted by Faul and Packer (2015), the contribution of the HLF remains something of a question mark – with no evidence that the Forum was never convened. Conversely, GEMs were assembled on two separate occasions (2012 and 2014), although their impact or ability to provide strategic guidance for the EFA project also remains unclear (EFA-GMR/UNESCO, 2015).

In practice, the leadership of the EFA architecture ended up falling on the Steering Committee – a body created in 2012 to replace a (largely inactive) International Advisory Group[22] and expected to "provide strategic guidance on all aspects of EFA, i.e., monitoring, research, global advocacy, knowledge-sharing and partnerships for specific issues such as financing" (EFA-GMR/UNESCO, 2015). The EFA Steering Committee became the main *locus* for reaching political consensus and the single nodal point bringing together the myriad of organizations involved in the "realization" of the EFA promise – that is, the original EFA convening agencies along with a selection of country representatives, the GPE, representatives from the civil society and NGO sector, the OECD and the GPE (Yamada, 2016).

Within UNESCO's Education Sector, the allocation of responsibilities for the coordination of EFA followed a similarly erratic course. Immediately after Dakar, the EFA mandate was transferred to an in-house informal coordination mechanism led by the interim Assistant Director-General for Education (ADG-E), with which different officials collaborated on an ad hoc basis. This group later became the Dakar Follow-up Unit, integrated within the Executive Office of the Education Sector and which operated as a

---

[21] Specifically, the Framework for Mutual Understanding (2001), the International Strategy (2002) and the Global Action Plan (2005/06).

[22] The International Advisory Group had been created as a mechanism for inter-agency cooperation, and was convened on two occasions, 2007 and 2008 (Robinson, 2014).

separate division. In the mid-2000s, the team became a section of a different division, being subsequently placed under the supervision of the ADG-E – and only at this point did it become the EFA Coordination Team. It should be noted that the team has tended to be under-staffed and has experienced high levels of turnover in its leadership positions – remarkably lacking a director between 2012 and 2014 (UNESCO Internal Oversight Service, 2016).

Overall, the EFA architecture in place for the period 2000-2015 had a remarkably unstable character – following an all too familiar pattern that echoed the difficulties faced by the EFA Forum during the post-Jomtien decade. An ever-changing division of labor, compounded by the imperative to accommodate a growing diversity of EFA constituencies, prevented the development of effective coordination mechanisms and translated into the further erosion of UNESCO's leadership capacity[23]. Different commentators noted that UNESCO had not been able to capitalize on its putative role as EFA lead – in fact, in its 2015 assessment of the EFA movement, the Global Education Monitoring Report (GEMR) reached the unflattering conclusion that:

> Overall, the formal EFA coordination mechanism, led by UNESCO, did not rise to the challenge of ensuring continuous political commitment and had limited success in actively engaging other convening agencies and key stakeholders. Many of the most successful mechanisms, initiatives and campaigns reviewed in this section happened *in spite of, rather than as a result* of, attempts at global coordination. (EFA-GMR/UNESCO, 2015, p. 41; emphasis mine)

*2.3.2. Monitoring EFA*

The very limited success of coordination efforts contrasts greatly with the progress made in terms of monitoring the achievement of the EFA goals. In particular, the establishment of the Global Monitoring Report (GMR[24]) has been repeatedly judged as a success by a variety of observers as well as by three independent evaluations (see for instance Ipsos Mori, 2018; or Universalia Management Group, 2010).

The early attempts at the production of an authoritative report had rather limited success. As reported by Packer (2007), the first report prepared in 2001 by the Dakar Follow-up

---

[23] As discussed in Chapter 5, with the passage of time the EFA Steering Committee tended to turn its attention to the definition of what came to be known as the post-2015 agenda. Remarkably, it proved far more effective in this function than in its coordination role. This is consistent with Faul and Packer's observation that "UNESCO-led mechanisms appear to be more comfortable and productive when focusing on the organization of global conferences (Dakar 2000; Incheon 2015), managing EFA national and regional assessments […], and in helping to define frameworks (as through the post-2015 Global Consultation with UNICEF" (p. 30).

[24] While the GMR is currently known as the Global Education Monitoring Report (GEMR), this chapter refers generally to GMR as the title of the report only changed with the adoption of the SDG agenda.

Unit was met with a tepid response from the HLG – with some of its members calling for a more robust, sophisticated product. It was against this background that an ad hoc, specialized team was created in a meeting facilitated by UNESCO early in 2002. The team was to work closely with the UNESCO Institute for Statistics (UIS), who had put in place a dedicated EFA Observatory (see section 2.3.3 below). It should be noted that the creation of the GMR did not occur without friction. This was as a consequence of the limited trust placed in UNESCO by the rest of the EFA sponsors, and the reluctance of some UNESCO officials. As documented by Daniel (2010), the GMR was received with little enthusiasm by part of the UNESCO Education Sector – especially those whose work was less clearly connected with the EFA agenda, and who resented the fact that UNESCO was "trading" the World Education Report collection (one of UNESCO's flagship product at the time[25]) for a publication with a much narrower view.

Partially in response to the limited credibility enjoyed by UNESCO in the wake of Dakar, the GMR team was afforded a considerable degree of independence from UNESCO – in both editorial and operational terms. Thus, the GMR publication is not conditional on the approval of UNESCO's Director-General, and even if the GMR is part of the UNESCO bureaucratic and administrative apparatus, it has, since its inception, enjoyed a considerable degree of managerial autonomy (for instance in the use of funds or hiring practices). The GMR has also been granted considerable financial independence from UNESCO. During the early 2000s, a series of donors committed to secure a sufficient level of funding for the production of the report, on the understanding that the absence of a rigorous analysis of the progress towards EFA was at the root of the limited traction enjoyed by the agenda set in Jomtien. Thus, the GMR was to be financed with extra-budgetary resources, primarily, through the contributions of a limited group of bilateral agencies[26]. As noted by a range of commentators (cf. King, 2010), the new team seized the opportunity afforded by these privileged circumstances, and managed to establish the GMR as a highly-regarded and widely-circulated compendium (Packer, 2007)[27] – a reputation that it has been able to preserve despite the emergence of other international education reports (Ipsos Mori, 2018).

---

[25] See below for an extended comment.

[26] It should be noted that the pool of donors has tended to diversify and expand with the passage of time. Such dynamics cannot be dissociated from the fact that, despite its many advantages, GMR's reliance on external donors has also translated into considerable financial instability (Ipsos Mori, 2018).

[27] The influence of the GMR is particularly well-established among development agencies, international NGOs and also some local NGOs and CSOs in low-income countries. Conversely, the GMR has been less intensively used by the EFA coordination bodies (Packer, 2007). Similarly, its influence on policy-makers in the Global North is far more limited since the report continues to be perceived as focusing on developing countries, and given the existence of comparative publications and data digests that focus on high-income countries, most notably those prepared by the OECD and the European Union (Ipsos Mori, 2018).

Overall, there is a consensus that the GMR succeeded in filling a vacuum and in establishing itself as a credible, authoritative source. A detailed comment on the dynamics behind such a rise is provided by Edwards, Okitsu, da Costa and Kitamura (2017). The authors argue that the solid reputation and authority enjoyed by the GMR stems from three main factors – the reliance on credible statistics, the quality of the production process (informed by an extensive consultation process and the commissioning of a number of background papers), and the clarity and relevance of its analytical work. In addition, the GMR has greatly benefited by its independence vis-à-vis UNESCO. Not only has the GMR remained unaffected by UNESCO's funding limitations but, more importantly, it has succeeded in removing itself from UNESCO's internal politics. The GMR has thus been able to engage in a critical assessment of country progress without interference of UNESCO's Board, as well as to offer credible comment on UNESCO's performance as coordinator of the EFA agenda. According to Edwards and colleagues (2017), the reputation enjoyed by the GMR has even resulted in a partial improvement of UNESCO's credibility, as the report is perceived by many as a UNESCO in-house publication – and UNESCO itself has been eager to promote it as one of its flagship products.

### 2.3.3. A fresh mandate for UIS

Another area in which Dakar acted as a catalyst was that of data gathering. As discussed above, quantitative and time-bound targets had been a conspicuously absence in the Jomtien agenda. While their exclusion had been a necessary move in order to secure consensus, this soon proved problematic for it rendered particularly challenging any effort to monitor progress towards EFA. In view of such difficulties, a decision was made in Amman 1996 to strengthen data-collection efforts in preparation for the end-of-decade assessment. An agreement was reached to monitor 18 indicators covering the six Jomtien dimensions – even if most of them in fact focused on primary education (Roberts, 2005; Torres, 2000). UNESCO was tasked with the coordination of inter-agency efforts in data gathering and particularly with the harmonization of the data submitted by countries in the form of national reports (Ito, 2012).

It needs to be born in mind that, by the late 1990s, UNESCO was already trying to actively enhance its own reputation in the data-gathering area. To this purpose, it engaged in a thorough reform of its statistical services, which crystallized in the creation of the UNESCO Institute for Statistics (UIS) in 1999 (Cussó, 2006). The key role assigned to UNESCO in the preparation of the 2000 End-of-Decade Assessment was thus effectively seized by a newly created UIS as an opportunity to raise its organizational profile. Not coincidentally, global education targets have since remained one of the UIS core areas of focus. Whilst the strategic reliance of the UIS on EFA (and later on SDG4) is the focus of Chapter 6, it is important to stress here that Dakar was one of the first opportunities for UNESCO to reassert its mandate in the area of data collection and to revitalize its lead in the production of statistics after a decade of decline.

UNESCO had started its work on cross-national statistics in the 1950s – crucially, with the support of the World Bank (Mundy, 1998) and largely as a consequence of the US insistence on the centrality of statistics work within the UNESCO mandate (Jones with Coleman, 2005). The UNESCO Statistical Yearbook, launched in 1963, soon became a trusted source of cross-national statistics and, for a long time, UNESCO remained the only international organization regularly producing comparative education statistics with a global coverage (Cussó, 2006). However, during the 1980s the relevance and reputation of UNESCO's statistical work started to deteriorate. With most international organizationsm and governments increasingly concerned or interested in students' skills, the relationship between education and economic competitiveness and budget efficiency, UNESCO's narrow focus on schooling and literacy data proved increasingly unsatisfactory (Cussó and D'Amico, 2005). In addition, the very reliability of UNESCO's statistics started to be called into question – to the point that, in 1993, a UNICEF-World Bank joint study concluded that UNESCO statistics were not only excessively narrow in scope, but also unreliable and insufficiently accessible (Puryear, 1995). A similar conclusion was echoed by Heyneman's notorious essay *The sad story of UNESCO's education statistics* (1999). By the late 1990s, UNESCO had been effectively displaced by the OECD[28] as one of the main providers of international education statistics, while UNICEF and Eurostat were starting to produce figures of their own (Cussó, 2006). The competitive dynamics in which these organizations engaged led to a situation of statistical fragmentation without precedent – one in which the harmonization efforts performed by UNESCO became an increasingly challenging exercise (Cussó and D'Amico, 2005).

To be sure, and partially in response to the dynamics initiated in the 1980s, a series of efforts had been made since the early 1990s in order to avoid a further erosion of UNESCO's status in the realm of international statistics. In particular, a UNESCO-OECD joint initiative had been set up in 1997 – the World Education Indicators (WEI) project, financed by the World Bank and oriented towards producing comparable statistics for a selection of OECD and non-OECD countries (cf. Grek, 2020). In response to the countries' growing concern over the performance of national education systems, in 1992, UNESCO had launched a project to monitor student achievement in selected countries, the Monitoring Learning Achievement initiative, implemented jointly with UNICEF. With a similar objective, the Southern African Consortium for Monitoring Education Quality (SACMEQ) had been launched in 1995 with the support of IIEP-UNESCO, and tasked with monitoring student learning through the production of comparable reports (Cussó and D'Amico, 2005). Finally, the launch of the World Education Report (WER), published during the 1990s on a biennial basis, also had a restorative purpose. As noted

---

[28] To be sure, the organizational overlap between UNESCO and the OECD in relation to data collection and the production of education indicators, dates back to the late 1960s (Ydesen and Grek, 2019). However, because UNESCO remained the most comprehensive source of cross-national education data, and it managed to preserve an image of technical robustness well into the 1980s (Heyneman, 1999), the two organizations had been competing on an equal footing until that moment.

by Mundy (1999), this was the first compendium-like product comparable to those produced by the World Bank or UNICEF. Despite the limited resources devoted to the project, the publication succeeded in boosting UNESCO's credibility in the realm of data-gathering and analysis. The significance of the WER is also synthesized by Jones, who notes:

> Successive issues of the WER indeed raised UNESCO's professional standing, especially through the careful responses made to prevailing World Bank and "Washington consensus" patterns of educational analysis and policy prescription, and through the considerable advances being made in the quality and application of UNESCO's statistical capacity. (Jones with Coleman, 2005, p. 68-69)

The mounting criticism faced by UNESCO's Division of Statistics eventually led to a restructuring of the UNESCO statistical program, a process well chronicled by the work of Cussó (2006) and Heyneman (1999). With the aim of identifying the most urgent areas for improvement and possible courses of action, a report was commissioned to the Board of International Comparative Studies in Education (BICSE)[29], a forum integrated within the American National Research Council with the financial support of the World Bank and two agencies of the US government. The final report entitled *The Worldwide Education Statistics: Enhancing UNESCO's Role* was published in 1995 and had a profound impact on UNESCO's thinking, decisively shaping the development of its statistical program in the years to come. The BICSE report did not focus so much on quality issues, but stressed the need for a change in the very orientation of UNESCO's statistical service, encouraging a function of policy orientation more attentive to student performance issues and on the relationship between education and economic competitiveness. It also advocated for an increase in the managerial autonomy of UNESCO statistical services.

It is in this report that the idea of a semi-autonomous specialized institute appeared for the first time. In 1997 the UNESCO Secretariat brought this idea before the General Conference for consideration and, after its approval by Member States, a rapid and highly contentious process of transition began. The design of the new institute was primarily informed by the so-called Thompson Report. Commissioned by UNESCO to a private firm, the report advanced a series of recommendations in relation to the future Institute's staff policy, structure and functions – most of them entailing a radical departure from UNESCO's uses. The UIS was thus established in 1999 amidst a climate of tension and discontent within UNESCO. As noted by Cussó (2006), the transition needs to be understood as a political exercise oriented at signaling the Institute's autonomy vis-à-vis UNESCO – as a means to appease some of its most prominent critics and to attract support from other multilaterals (most notably, the World Bank) and high-income governments. This is why UIS's managerial, political and financial autonomy were maximized despite the unease caused by such changes among UNESCO's personnel. Such commitment to

---

[29] The Board had been set up by the US National Research Council in the late 1980s in order to examine the comparability of education statistics.

institutional independence was consolidated by the decision to locate the UIS outside UNESCO's headquarters (specifically, in Montreal), as well as by the almost total re-staffing of the Institute in the early 2000s. Conversely, efforts to enhance the quality and robustness of the produced data were much less prominent during the restructuring of UNESCO's statistics services. In fact, the first years of existence of the UIS were characterized by a conspicuous lack of progress in the area of data management[30], and by the slowdown of activities relating to data-processing, verification and analysis. It is also remarkable that one of the few recommendations advanced by the Thompson Report in relation to data production concerned the harmonization of the UIS methods with the OECD guidelines – a development that confirmed the growing influence and dominance exerted by OECD in the field of international education statistics.

While the evolution of the UIS and its critical transitioning to the "post-2015" era are not the subject of this chapter, suffice to say that by the time of WEF 2000, the position of the UIS was in many senses a precarious one. Its mandate and rapport with UNESCO remained indefinite and contentious, while its added value in relation to the work performed by other international organizatiosn (and the OECD in particular) was declining. The heavy involvement of the UIS with the EFA End-of-Decade Assessment and the monitoring of the EFA goals needs to be understood against a specific background – and especially in relation to the need for the UIS to conform to the different expectations placed on the Institute by UNESCO, by donors, and by governments. EFA was thus approached as an opportunity for the revitalization of UNESCO's statistical services, in an attempt to enhance both their internal and external legitimacy.

### 2.3.4. The creation of the FTI

Another of the key transformations spurred by the WEF was the creation of the Fast Track Initiative (FTI). Conceived as a mechanism to scale up, pool and (re)distribute education aid towards those countries most in need, the initiative was largely motivated by the Dakar commitment that "no countries seriously committed to education for all will be thwarted in their achievement of this goal by a lack of resources" (World Education Forum, 2000, paragraph 10)[31]. The FTI was formally launched in 2002 by the President of the World Bank (James D. Wolfensohn) in a high-profile event attended by a number of political personalities. Even if there was considerable uncertainty surrounding the project (for the details of the initiative had not yet been worked out at the time of the announcement), it was trumpeted as "an historic first step towards putting all developing

---

[30] Particularly in relation to the maintenance of databases (automation of calculations, updating of demographic series, revision of projections, etc.).

[31] Despite such commitment to the EFA goals, the FTI agenda was for a long time essentially driven by the MDGs, with much of its activity focusing on the promotion of UPE. As discussed later, this led to tension between the FTI and other UN agencies and the NGO sector, who advocated for a broader vision more clearly aligned to the EFA agenda (Bermingham, 2011).

countries on an education Fast Track that could transform their social and economic prospect" (BBC, 2002, as cited in Bermingham, 2011, p. 559).

The establishment of the FTI soon became a source of major disruption – arguably a predictable outcome in a field fraught with competition and a general lack of trust among the main education agencies. In fact, the very creation of such a mechanism had been advanced by the World Bank as a means to avoid relying on the UNESCO-led High-Level Group for the purposes of aid delivery and financial mobilization (Faul and Packer, 2015). The multiple tensions that shaped the FTI in its early years as a result of both inter- and intra-organizational competition have been discussed at length in Bermingham's (2009, 2011) insider accounts[32]. In relation to intra-organizational tensions, the author observes that the FTI became a source of disagreement *within* the World Bank, as the new initiative was met with reluctance by the World Bank education sector group. This was so as, from the start, budget support programs had been identified, as the World Bank financing modality more clearly aligned to the FTI objectives. Since this modality had typically been managed by country economists or other financial specialists, the education sector group feared that the FTI would accelerate the downsizing of specialist education staff. In order to address such concerns and reconcile them with the imperative to promote country ownership, the wording of the FTI Framework (i.e., the FTI "unofficial constitution" during its first years of operation) was carefully negotiated so that it expressed the priority given to "the use of country systems" without mentioning a specific aid modality (Bermingham, 2009). However, the reliance on budget support operations remained an exception by the late 2000s, as support was typically delivered through arrangements similar to those of education investment projects (Bashir, 2009).

When it comes to inter-organizational tensions, these were largely caused by the prominent character of the World Bank in the management and administration of the FTI. It needs to be born in mind that the FTI Secretariat was initially established as a team within the World Bank's Human Development Network, being physically located within the World Bank's headquarters and with the totality of the staff being classified as World Bank employees. Such dependence became increasingly problematic as the FTI expanded and as the Secretariat took charge of a growing number of responsibilities (Bermingham, 2011). This lack of institutional distance was compounded by the multiple roles played by the World Bank – who by the mid-2000s acted simultaneously as the supervising agency for most funding operations, the implementing agency for the associated education plans, and a member of most local donor groups, *i.a.* Such arrangements created a number of conflicts of interest and led to confusion at the country level; the FTI was frequently perceived as a World Bank program (Bermingham, 2009). In addition, and as reported by Menashy (2017), the prominence of the World Bank in the FTI management was seen by some as an opportunity for the World Bank to advance its own policy preferences. The disproportionate influence wielded by the World Bank was in

---

[32] The author was formerly the Head of the FTI Secretariat and has acted as the Co-Chair of the FTI Steering Committee.

turn exacerbated by the limited clout of the FTI Steering Committee (FTI-SC) and the fact that this body was long dominated by donors. The limited involvement of UNESCO and UNICEF was particularly striking in this regard[33], as was the limited representation of the NGO and developing countries (Bermingham, 2011).

Such dynamics became the target of growing criticism and were explicitly highlighted by the mid-term external evaluation of the fund (Cambridge Education/Mokoro/Oxford Policy Management, 2010) which eventually led to the restructuring and rebranding of the FTI. Since 2011 the organization has been rebranded as the Global Partnership for Education (GPE), and since 2009 it has been governed by a more diverse Board of Directors (an organism that replaced the FTI-SC). Also, efforts have since been made to demarcate clearer boundaries between the GPE and the World Bank. The process has been discussed extensively by Menashy (2017), who concludes that, despite efforts to turn the organization into a more participatory and democratic space, power asymmetries persist, and the World Bank retains an unbalanced influence over GPE operations[34].

It should be noted that *governance issues* were not the only source of criticism documented during the founding years of the FTI. Other relevant areas for concern included the delay in the disbursement of certain funds. This was largely as a consequence of the 2007 decision to apply standard project procedures (including lengthy financial management and safeguards assessment) to all operations managed by the World Bank, in an attempt to appease concerns over financial security expressed by World Bank senior officials (Bermingham, 2011). Another recurrent area of concern was the reliance on indicative benchmarks which operate in practice as targets against which country progress is assessed. A prominent critic of the problematic character of such arrangement was Rose (2005), who argued that the use of such benchmarks risked advancing de-contextualized policy options that do not sufficiently take into account the political and economic viability of the proposed national education plans. Finally, another longstanding source of debate has been the (arguably limited) alignment of the FTI with the principles of country ownership, aid effectiveness and donor harmonization advanced by a number of declarations during the early 2000s – most notably, the Monterrey Consensus, the Rome Declaration and the Paris Declaration (cf. Bermingham, 2009; Menashy, 2017).

For the purposes of this dissertation, the creation of the FTI remains a crucial episode since, with the passage of time, the organization has tended to overshadow the EFA Steering Committee as "the forum of choice" among a wide range of global education policy actors (Faul and Packer, 2015). A similar point is raised by Tikly (2017) in his recent overview of power dynamics within the "EFA regime". According to the author, the GPE would have ended up morphing into an alternative center of gravity within the

---

[33] Hence, UNESCO only joined the FTI-SC in 2004 but was not included in the trust fund meetings where key decisions were being made. UNICEF only became a member in 2006 (Bermingham, 2011).

[34] See Mundy (2016) for a synthetic overview of recent transformations.

EFA community, thus contributing to the further erosion of UNESCO's leadership capacity.


*2.3.5. The repositioning of NGOs as a policy interlocutor*

Finally, the WEF also brought about a series of changes in relation to the role and position of the international non-governmental organizations (INGOs) within the education realm. More specifically, it gave impetus to the articulation of the Global Campaign for Education (GCE) – a multi-NGO platform that greatly enhanced the visibility, presence and authority of the NGO sector in global spaces of education governance.

As reported by Mundy and Murphy (2001), while transnational advocacy networks had been growing in importance at least since the 1970s, in the field of education this had been a comparatively late phenomenon since the NGO sector had remained remarkably fragmented during the 1980s. In addition, the interlinkages between INGOs and IGOs had long remained highly superficial, with the consequent absence of the former from IO-sponsored conferences and meetings. While the 1990s marked a turn away from such dynamics (with an explosion of INGO-activity explicitly focusing on education and the parallel recognition of their role on the part of IOs), the influence of the NGO sector in the area of norm-building and goal-setting remained limited.

As noted above, while the WCEFA constituted one the first international meetings with a significant representation of NGO participants, the impact of this constituency was limited. This stemmed partially from the fact that, as reported by Chabbot (1998), an explicit effort was made on the part of the Secretariat to prevent NGOs forming a cohesive group. In addition, the NGO sector was incorporated essentially in the light of its role as provider of education, or as a key collaborator in the delivery of education. Such a conception had a direct impact in the selection of organizations invited to the conference. This selection privileged large Western-based institutions that focused on development and relief activity, but left aside a number of NGOs that focused on adult education and literacy (Mundy and Murphy, 2001).

The lead-up to Dakar WEF marked the end of these dynamics. In 1999, three NGOs (Oxfam, ActionAid, and Education International) launched independent campaigns focusing on the EFA commitments. That same year, they teamed up to launch the Global Campaign for Education, explicitly targeting the review of EFA progress, along with the UN "follow-up conference" of the Social Summit and the G7 Spring Meetings. Later on, the group was joined by the Global March against Child Labor as well as other Southern-based coalitions, which contributed to make the campaign more inclusive and less Northern-oriented. In the months leading to the Dakar meeting, the GCE gained considerable visibility through the organization of a number of activities, including press conferences and meetings with national delegations. This enabled the GCE to overcome attempts to limit NGO presence on the part of the WEF organizers. As documented by

Mundy and Murphy (2001), the initial efforts to keep a tight limit on the number of NGOs attending the WEF were reversed the day before the start of the forum. This signaled a clear move towards the recognition of NGOs as policy interlocutors[35].

Such a development was partially afforded by the greater coordination and cohesiveness of the NGO sector[36]. Shortly before the Dakar conference, GCE members adopted a common list of demands (the so-called bottom-line platform) with the purpose of using it as their basis for lobbying efforts during the forum. In addition, during the conference they strategized together in the election of NGO representatives in the Drafting Committee and the Futures Committee. The strategy proved successful for, as documented by Mundy and Murphy (2001), the input of the GCE decisively shaped a number of decisions and commitments made at the conference – most notably, the "inclusion of the wording 'free' education; endorsement of the idea of national educational forums and an expanded definition of education that includes commitment to early childhood education and adult literacy; and a commitment to annual high-level EFA review meetings" (p. 121). In addition, the very creation of the FTI cannot be dissociated from GCE's calls to increase bilateral and multilateral aid for education, and the proposal to establish an international funding mechanism for the delivery of educational assistance.

However, the impact of the GCE did not end with the Dakar forum. On the contrary, the conference marked the start of a new era in the relationship between NGOs and IOs, and changed the way in which the non-governmental sector inserted itself in global governance spaces. Such developments have been examined by Mundy (2012) and Gaventa and Mayo (2010), who have discussed both the evolution of the GCE and their achievements during the 2000s[37]. As reported by these authors, the GCE succeeded in preserving a strong presence at a number of global forums, especially within the EFA architecture. It thus gained access to the EFA High-Level Working Group, the FTI Board and the GMR Board. According to Mundy (2012), part of the success of the GCE stems from its ability to play "an insider/outside role"– that is, engaging directly in such spaces

---

[35] As noted above, the participation and influence of the NGO sector had been rather limited in previous conferences as well as during the follow-up of the EFA. This had however started to change with the incorporation of Oxfam in the EFA Forum Steering Committee. This NGO advocated for an ambitious Global Action Plan (GAP) and especially for the increase in the financial resources available for the EFA agenda. As noted by Little and Miller (2000), the GAP proposal along with the presence of Oxfam in the Steering Committee led to "the greatest divergence of views among Steering Committee members since its inception" (p. 40). Such development is indicative of the growing legitimacy and leverage exerted by the INGOs in the structures and spaces of global education governance.

[36] Even if, as noted by Mundy and Murphy (2001), it is important not to overstate the cohesiveness and unity displayed by this group, a number of divergences and fault lines emerged during the forum, not only between the GCE and other NGOs, but also within the GCE coalition.

[37] This section addresses the impact of GCE at the global level. It does not grapple with the role and influence exerted by GCE national coalitions, or the regional networks. A number of case studies focusing on the role of national coalitions can be found in the edited volume *Campaigning for "Education for All"* (Verger and Novelli, 2012).

while being critical of them or of their performance. Thus, through a combination of insider techniques (i.e., collaboration with key decision-makers) and outsider strategies (i.e., more confrontational advocacy campaigns), GCE has succeed in decisively shaping a number of global debates relative to the implementation of EFA (and, as discussed in Chapter 5, its transitioning into the post-2015 era).

It needs to be noted that GCE's membership expanded considerably within a relatively short time. This was largely the consequence of the creation of a number of national coalitions themselves encompassing a varying number of local non-governmental organizations (NGOs), civil society organizations (CSOs), teacher organizations, parent associations, and so on. According to the most recent available count including a breakdown by membership category, GCE encompassed almost 100 national coalitions, 13 regional networks, and 14 INGOs (GCE, 2018). In addition, the volume of resources available to GCE also grew substantially on account of a number of successful fundraising efforts. As documented by Mundy (2012), at least during its first decade of existence, GCE proved itself skillful in the attraction of short-term funding for specific projects (such as the Real World Strategies program supported by the Dutch government, the 1Goal campaign funded by FIFA and a range of bilateral donors, the Civil Society Education Fund, and the Commonwealth Education Fund supported by the British Government).

While the progressive institutionalization of GCE is beyond the scope of this chapter, it should be noted that this process was not free from tensions which directly shaped GCE's engagement with global frameworks and global conferences in the years to come. Such tensions included struggles over the purpose of the campaign, in particular the trade-off between the EFA agenda and the MDGs. There were differences over the prioritization and framing of certain themes, in particular the risk of imposing "a Global North agenda", and there were debates on the representation of Southern voices in decision-making spaces. Differences even arose over the very identity and structure of the Campaign, with some members expressing a preference for a networked style close to that of social movements, whilst others advocated and advanced a more hierarchical, formalized style characteristic of international federations (Gaventa and Mayo, 2010; Mundy, 2012). Similarly, the heterogeneity exhibited by national coalitions in terms of advocacy and monitoring capacity, and their adherence to and engagement with GCE's global program, also challenged the very credibility of the Campaign, which translated into a rather unequal record of success. Finally, GCE's reliance on project-specific funding translated into a situation of financial uncertainty or unpredictability that has hampered efforts to consolidate a stable governing structure (Mundy, 2012; Tota, 2014).

## 2.4. Advancing education through the Millennium Development Goals

Less than six months after Dakar, the Millennium Declaration (MD) was approved in the context of the Millennium Summit held in September 2000. This text would eventually

give rise to the Millennium Development Goals (MDGs), announced in 2001. With two of the MDGs directly addressing education (see Table 2.4 below), this new set of international goals soon eclipsed the EFA agenda within the international development community[38]. This section addresses the origins of the MDG framework in connection with prior exercises of goal-setting sponsored by the UN. It also examines the iterative process through which the goals were negotiated, identifying the main ideational and organizational forces that shaped their final form.

**Table 2.4. Millennium Development Goals**

| Goal | Associated target |
|---|---|
| *MDG 1 – Eradicate extreme poverty and hunger* | 1 - Halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day.<br><br>2 - Halve, between 1990 and 2015, the proportion of people who suffer from hunger. |
| *MDG 2 - Achieve universal primary education* | 3 - Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling. |
| *MDG 3 - Promote gender equality and empower women* | 4 - Eliminate gender disparity in primary and secondary education, preferably by 2005, and in all levels of education no later than 2015. |
| *MDG 4 - Reduce child mortality* | 5 - Reduce by two-thirds, between 1990 and 2015, the under-five mortality rate. |
| *MDG 5 - Improve maternal health* | 6 - Reduce by three quarters, between 1990 and 2015, the maternal mortality ratio. |
| *MDG 6 - Combat HIV/AIDS, malaria and other diseases* | 7 - Have halted by 2015 and begun to reverse the spread of HIV/AIDS.<br><br>8 - Have halted by 2015 and begun to reverse the incidence of malaria and other major disease. |
| *MDG 7 - Ensure environmental sustainability* | 9 - Integrate the principles of sustainable development into country policies and programmes and reverse the loss of environmental resources.<br><br>10 - Halve, by 2015, the proportion of people without sustainable access to safe drinking water and basic sanitation.<br><br>11 - By 2020, to have achieved a significant improvement in the lives of at least 100 million slum dwellers. |
| *MDG 8 - Develop a global partnership for development* | 12 - Develop further an open, rule-based, predictable, non-discriminatory trading and financial system.<br><br>13 - Address the special needs of the least developed countries.<br><br>14 - Address the special needs of landlocked developing countries and small island developing States.<br><br>15 - Deal comprehensively with the debt problems of developing countries through national and international measures in order to make debt sustainable in the long term. |

---

[38] It is remarkable that even organizations whose legitimacy stems directly from the EFA mandate have focused the bulk of their operations on the attainment of MDGs – FTI is a case in point (Packer, 2007). Similar dynamics are reported by bilateral agencies such as DfID, the World Bank or UNDP  (Colclough, 2005; King and Rose, 2005). The tensions generated by the coexistence of these two sets of goals are discussed in greater detail in Chapter 5.

Source: Adapted from UNGA (2001).


### 2.4.1 The revitalization of global conferencing

In order to understand both the origins and significance of the MD and the MDGs, it is necessary to take a step back and situate these exercises in relation to a sequence of UN thematic conferences celebrated during the 1990s. Thus, the Millennium Summit was explicitly devised as a stock-taking exercise and as a culmination of this cycle of subject-specific conferences.

This specific modality of soft summitry had been reactivated in 1990 after a decade of deceleration and general skepticism regarding the potential of large gatherings. Up to seven major UN conferences[39] were held over a five-year period, namely, the World Conference on Education for All (Jomtien, 1990), the World Summit for Children (New York, 1990), the Earth Summit (Rio, 1992), the World Conference on Human Rights (Vienna, 1993), the International Conference on Population and Development (Cairo, 1994), the World Social Summit on Social Development (Copenhagen, 1995), and the World Conference on Women (Beijing, 1995). These conferences were partially justified as a follow-up to conferences on the same topic held during the 1970s, and they provided a range of IOs with a "fresh mandate" on their own area (Browne and Weiss, 2014).

For the purposes of this dissertation, the World Summit for Children deserves special mention for, as noted by Emmerji, Jolly and Weiss (2001), it played a key role in revitalizing the enthusiasm for UN conferences explicitly committed to goal- and target-setting, and became a blue-print for the conferences that followed. The success of the conference in securing political commitment and additional resources for the cause, as well as the robustness of the follow-up mechanisms it established (which, unlike Jomtien, succeeded in maintaining the momentum well beyond the conference), sparked a new wave of optimism regarding the potential of conferences and target-setting (Hulme, 2010). Other high-profile conferences such as the UN Conference on Environment and Development (commonly known as Río Summit or Earth Summit) had a similar effect, rendering evident the potential of such encounters as vehicles to mobilize public attention even when they failed to secure a political consensus.

---

[39] Note that, in these events, the term summit is used in a loose, popular sense – officially speaking, most of these events were not UN summits but world summits sometimes held at the UN – as in the case of the Children's Summit.

As discussed above, UN conferences had experimented a rapid process of institutionalization during the 1970s. This process continued throughout the 1990s – most of the thematic conferences[40] celebrated during this decade indeed shared a number of common attributes (Emmerji et al., 2001; Haas, 2002). This process of progressive homogenization was in fact recognized early on by a number of observers. In a 1996 essay, Fomerand noted that even if it was not possible to identify an "ideal type" of UN conference (as they drew on a variety of organizational and planning models), they all relied on a number of established practices in relation to the selection of participants, the proceedings, and the products or outcomes (Fomerand, 1996). The basic format of such encounters was also synthesized by Clark, Friedman and Hochstetler (1998):

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

Other central features of such events are highlighted by Hulme (2010), who draws attention to the key role of ad hoc secretariats (in charge of both the convening of preparatory meetings and the consultations with experts, NGOs and so on). Hulme also highlights the ordinary establishment of follow-up meetings to review progress – typically, five and/or 10 years after the conference.

## 2.4.2. IDGs, a stock-taking exercise

1995 is described by Hulme (2010) as the "peak year for summitry", with the celebration of two major conferences (the World Summit for Social Development in Copenhagen and the UN Fourth World Conference on Women) separated by only six months. After this, a period of "conference fatigue" ensued – that is, a sense of exhaustion and overload in view of the proliferation of recommendations, priorities and goals covering an excessively wide range of topics. Even if global summitry did not stop altogether after 1995, it certainly slowed down.

---

[40] This includes conferences understood as one-off events, called on a non-routine basis and which focused on a single-issue area deemed to be of international importance. As noted by Haas (2002), conferences need to be distinguished from summits which are generally organized on an ongoing basis, attended by a small group of ministers and top-level bureaucrats and are less open. Conversely, "UN conferences enjoy greater and broader political legitimacy by virtue of their universal representation and the opportunity for middle-level powers and small powers to have a say" (p. 77).

It is against this background that the OECD's DAC decided to engage in a synthesis exercise oriented at mobilizing support for foreign aid and assistance within OECD countries and, more importantly, at providing donors with a sense of direction (Fukuda-Parr and Hulme, 2009). The idea first emerged in the context of the DAC High-Level Meeting of 1995, and was largely envisaged as a response to the declining levels of ODA. This trend had become a widely shared concern within OECD's ministers of international development (or top bureaucrats within aid agencies), who feared becoming increasingly irrelevant in the context of their respective national administrations. With the purpose of countering such tendencies and of reflecting on the future of ODA, a *Groupe de Réflexion* was created. It is in the context of this Groupe that the idea of a "listing exercise" emerged, oriented towards making sense of a decade of UN summit declarations. As noted by Hulme (2007), this listing became the focus for the Groupe in a rather unexpected way. The first listing was prepared by the DAC staff essentially as a background document. However, the list soon caught the eye of a number of development agencies, who saw the potential of a coherent list of targets as a management device. This was particularly the case for the US, UK, Germany and Norway delegations. As noted by Black and White (2004), the enthusiasm of these members cannot be dissociated from the growing prominence of Results-Based Management (RBM) in these countries. These were in fact countries that had recently begun to engage with a reform of parts of the public administration on the basis of RBM-related principles that put a premium on the use of targets and performance outcome measures.

The listing exercise eventually gave rise to the document *Shaping the 21st Century: The Contribution of Development Co-operation* – a 20-page publication containing a list of six "ambitious but realisable goals" (OECD, 1996, p. 2). Known as International Development Goals (IDGs)[41], these goals soon became the main focus of interest, not only among media outlets, but also among OECD governments. The goals attracted a good deal of rhetoric support on the part of these governments and were formally endorsed at a number of OECD ministerial meetings and G7 gatherings, although they only gained political traction among a limited number of donors characterized by their progressive orientation (Denmark, the Netherlands, Norway and Sweden) (Hulme, 2007).

*2.4.3. Crafting a historical consensus*

While the 1996 IDGs did not cause an immediate reaction within the UN system, they certainly informed the UN discourse and action around human development in the years to come. Partially as a consequence of this development, by the late 1990s the proliferation of goals started to be explicitly recognized as problematic within UN quarters. This became a personal preoccupation of the then Secretary-General Kofi

---

[41] Sometimes known as International Development Targets (IDTs), particularly within British circles. This preference seems to stem from DfID's reference to the IDTs instead of IDGs (Hulme, 2010).

Annan, who concluded that the profusion of goals conveyed a confusing message while rendering it impossible to identify specific priorities (Bissio, 2003; Fukuda-Parr and Hulme, 2009).

Such preoccupations and the influence of the IDGs became particularly evident during the preparation and planning for the Millennium Assembly of the United Nations, to be celebrated in September 2000. It is important to bear in mind that the stakes associated with such an encounter were high – the UN's 50th Anniversary Summit celebrated three years before had been judged a failure. The Millennium Summit was thus approached as a means to enhance and restore the credibility of the UN (Hulme and Scott, 2010).

The planning process kicked off in May 1999, with Annan's identification of the main themes that should guide the Assembly and the celebration of a number of (formal and informal) meetings oriented at shaping the pre-summit report which, in turn, would inform the negotiation of the Millennium Declaration (MD). As discussed by Fukuda-Parr and Hulme (2009), such process was largely driven by concerns over communicative efficacy. The willingness to produce an actionable, coherent and simple declaration is central in explaining the particularities of the drafting process behind the production of the pre-summit report; this was highly atypical for UN standards. As noted by the authors, the drafting process for this kind of product is usually initiated by junior staff (typically from different departments) and is followed by a long process of negotiation during which UN departments, agencies and national delegations provide inputs to the document and lobby for the inclusion (or exclusion) of specific themes. While such processes are generally touted as highly inclusive and democratic, they are by definition lengthy and tend to give rise to distinctly verbose documents characterized by limited levels of clarity and internal coherence.

Since a consensus had emerged within the UN high command that this was an outcome to be avoided at all costs, the drafting process for the pre-summit report was firmly controlled – the saying goes that it was "drafted on the 38th floor" by John Ruggie and Andrew Mack, Annan's senior aides. Even if such accounts have been challenged as inaccurately reductionist (cf. McArthur, 2014), the process certainly entailed a departure from the more established and open modus operandi. The result was a product exhibiting a degree of coherence and clarity unusual for UN standards. The final report, entitled *We the Peoples: The Role of the United Nations in the 21st Century* and published in April 2000, was a turning point in that it signaled the UN willingness to change. However, the exercise did not succeed in superseding the agenda set by the OECD's DAC four years before. In comparison with the IDGs, the *We the Peoples* report did not fare well. While both documents were oriented towards setting the agenda for international development and poverty reduction, the UN report addressed a wider range of topics and diluted any sense of prioritization – largely as a consequence of the need to satisfy a broad range of constituencies, as per UN's conventions and usages (Hulme, 2007). In addition, they were less coherent with the RBM philosophy that oriented the IDGs.

Two months after the *We the Peoples* publication, the four major development multilaterals (the UN, OECD, the World Bank and the IMF) launched the document *A Better World for All: Towards the International Development Goals*. While the publication aimed at proving commitment and cooperation towards the same objective, it contributed to reveal the fragile position of the UN. This was because the document "re-iterated the DAC's 1996 IDGs almost exactly [*and*] it re-affirmed the primacy of results-based management thinking" (Hulme, 2007, p. 11). It also made it evident that global goal-setting was becoming a "twin-track process" – with the UN focusing on the Millennium Summit, and the OECD and the Bretton Wood institutions focusing on the IDGs (Hulme and Scott, 2010). The report received strong criticism from NGOs, CSOs and even some church-related organizations, who argued that the report represented a narrowing of the development agenda (Saith, 2006; Vandemoortele, 2011).

Against such an unpromising background, the drafting of the MD was again explicitly planned as a tightly controlled process, this time led by the General Assembly President and the Ambassadors of Guatemala and New Zealand. The strategy proved fruitful in that, as noted by Fukuda-Parr and Hulme (2009):

> The MD presents a relatively focused and powerful vision of what global poverty eradication might be. A more open process would have produced a "Christmas tree" with hundreds of goals, defeating the purpose of communicating a coherent global poverty eradication agenda. (p. 14)

This should not suggest that production of a consensual Declaration was a simple endeavor. As reported by Hulme (2007), those in charge of the process had to strike a delicate balance between a wide range of constituencies, each one of them exhibiting a distinct set of preferences. This included the UN member states and NGOs, social movements and private companies, organized in structured as well as informal coalitions, and frequently lobbying for a highly particularistic agenda. A wide range of consensus-crafting techniques had to be employed. Some of them were oriented towards making the very content of the Declaration more palatable to a wide range of agents. Thus, the potentially contentious issues (including the economic strategy necessary to achieve such goals) were deliberately left aside. In other words, the drafting process "took the edge off" the UN conferences and summits – avoiding reference to their most contentious proposals. Other compromise strategies were subtler and more formalistic in nature. This is, for instance, true of the division of the poverty reduction resolutions in two separate paragraphs, (one starting by "We resolve further", the other with "We also resolve") in order to signal different degrees of consensus (Hulme, 2007). A similar rationale lies behind the vagueness of the MD regarding the time frame over which the targets are to be achieved. Thus, even if the MD referred to 2015 as the *end year*, it never specified the *baseline*. The true level of ambition of goals such as "halving poverty by 2015" thus remained unclear. Such an ambiguous formulation aimed at reaching a compromise between developing countries. This called for a sense of realism, and for developed countries to be more favorable towards ambitious targets (Vandemoortele, 2011).

*2.4.4. From the Millennium Declaration to the MDGs*

The MD was essentially a normative document in which world leaders committed themselves to the consecution of a number of common objectives, but it lacked an implementation plan, and provided little detail on the reporting process. This was to be developed in a report of the UN Secretary-General (UN-SG) to the General Assembly – the so-called Road Map. Again, this was an unusual development for UN declarations and reports, driven by the communicative concerns of a segment of the UN bureaucracy (in particular the Secretary-General's Office). The top management of the World Bank and UNDP were also quick in recognizing the need for a clear set of targets. As noted by Fukuda-Parr and Hulme (2009), the individuals[42] that pushed for the "distillation" of the MD into a set of clear goals were driven by an organizational imperative rather than by a commitment to a particular set of ideas. In this sense, and as noted by the authors, they can be seen as message entrepreneurs rather than norm entrepreneurs. More specifically, the group aimed at bridging the sharp ideological divide that had emerged around the Washington Consensus, and at building more consensual relationships within the development community.

The "distillation" process, however, proved a challenging exercise. As noted above, the development community had been caught in a twin-track situation that pitted OECD's IDGs against UN's MD. As discussed by Hulme (2007), while reconciling the two sets of goals was perceived by many as the logical step, this was problematic for all four development multilaterals. The OECD was unwilling to abandon a device that it perceived as a major organizational asset, while the World Bank and the IMF had already endorsed the IDGs and had little reason to attach themselves to another set of priorities. The UN, in turn, could not afford to modify a political agreement (the MD) by merging it with a pre-existing set of targets. A final compromise was only reached through a division of labor between the Bretton Wood institutions and the UN. It was agreed that the World Bank and the IMF would support a list of goals and targets based on the MD if the UN gave the green light to their full authority over Poverty Reduction Strategy Papers (PRSPs).[43]

With this vision in mind, the task to develop the MDG was handed over to an ad hoc group led by Michael Doyle (UN-SG Office) and Mark Mallock Brown (UNDP), and composed of a selection of elite technocrats from the World Bank, OECD's DAC and UNDP (Eric Swanson, Brian Hammond and Jan Vandemoortele, respectively). Again,

---

[42] This group included Michael Doyle from the UN-SG Office, Mark Malloch Brown from UNDP, the World Bank's President James D. Wolfensohn, UK International Development Minister Clare Short and UN-SG Kofi Annan.

[43] Launched in 1999 as a result of a joint initiative by the World Bank and the IMF, PRSPs are national planning frameworks required for countries wishing to qualify for external assistance or debt relief. They represent a novel approach in that they combine social development goals (strategies to reduce poverty) with plans for economic growth and macroeconomic stability (Mundy, 2006).

the preference for these figures (instead of leaders of UN statistical offices) represented a move away from UN practices and was essentially oriented at preserving the coherence and simplicity of the final product. It was no coincidence that the group was a highly cohesive one – many of its members had collaborated before in the elaboration of *2000: A Better World for All* (Clegg, 2015; Fukuda-Parr and Hulme, 2009). According to Vandemoortele's (2011) inside story of the process, this composition was instrumental in fostering the process in that it facilitated the emergence of "a reasonable entente among the UN staff and those of the World Bank, IMF and OECD" (p. 7). In fact, and as noted by Hulme (2007), the final formulation of the MDGs drew largely on the IDGs as presented in the *A Better World for All* document – with the incorporation of some of the points advanced by the MD. This was so that the IDG proposal was perceived as more coherent and more clearly aligned to an RBM logic and, more importantly, it was the preferred option of most of the involved partners.

It should be noted however, that the IDGs were not reproduced faithfully – those elements judged likely to jeopardize the formation of a political consensus were excluded. This was the case of reproductive health, which was resisted by the Vatican and some Islamic states. It is also relevant that the MD declaration had a direct impact on the formulation of Goal 8 – the development of a global partnership for development. This was included in order to make the new list more palatable for developing countries, for it created a series of obligations for the Global North, including the provision of a percentage of GDP to aid. It is however remarkable that Goal 8 is the only proposal lacking a specific timeframe, and for which a number of its associated targets were not formulated in quantitative terms – thus being the single proposal not adhering to the RBM principles (Hulme and Scott, 2010).

Another of the strategies used to maneuver for compliance was the creation of a benchmarking system that tended to diffuse lines of responsibility. The conflict-avoiding dynamics behind this formulation has been extensively discussed by Clegg (2015), who notes that:

> Goals One to Seven were linked to precisely defined targets and indicators, but responsibility for the ultimate achievement of these outcomes was placed rather diffusely at the door of all members of the international system. Goal Eight specifically aimed to encourage developed countries to adjust their aid, trade, and other externally oriented policies so as to foster a more development enhancing environment; however, unlike the earlier elements of the framework, no quantitative metrics were tied to this aspect. The MDG authors were particularly wary of including the 0.7 per cent of GDP aid target, for fear of alienating in particular the US and Japanese leadership from the initiative. (p. 959)

*2.4.5. The MDG architecture*

The adoption[44] of the Road Map and the MDGs in September 2001 was followed by the creation of a dedicated bureaucracy within UN quarters[45] – most notably, the establishment of the Global Monitoring Reports[46] produced by the UN with the assistance of statistical staff from a variety of IOs. In addition, the UNDP underwent a number of transformations oriented at maintaining the momentum around the MDGs. The changes included the creation of the Millennium Campaign, oriented at the mobilization of political support, and the Millennium Project, a research-oriented project aiming at identifying the technologies and policy options necessary to accelerate progress towards the MDGs (Manning, 2009). In addition, UN Country Teams were encouraged to foster the integration of the MDGs into national planning and policy-making (Fukuda-Parr and Hulme, 2009). This institutionalization effort was subsequently strengthened by an advocacy strategy oriented at encouraging other development organizations (government or NGO, bilateral and multilateral) to commit resources and operations towards the realization of the MDGs. Such efforts crystallized in the 2005 UN World Summit, where governments resolved to adopt national strategies to achieve the MDGs (McArthur, 2014).

McArthur (2014) has noted that this multi-layered and intricate monitoring strategy reflects the tensions involved in the "global plus national" strategy pursued by the architects of the MDGs. The technical group in charge of the formulation of the goals had faced the decision of whether to establish a unifying list of targets or promote the use of national targets for monitoring purposes. The reliance on a common list of goals (i.e., the so-called "global plus national approach") was the preferred option of many of the experts in the group, but some of them (most notably, UNDP's Jan Vandemoortele) pushed for a "global but not national" approach as a means to protect developing countries from external pressures through the imposition of an agenda they could not possibly attain. While the 2001 Road Map containing the MDGs did privilege the "global plus national" approach, other documents that followed were characterized by a sense of ambiguity – leaving the door open for the adaptation of targets at the country level. This was true of the UN Development Group Guidance Note on country-level reporting, which noted that "The MDGs will need to be contextualized within the country-specific situation. In other words, the global goals and targets will need to be adapted into national equivalents with the endorsement of national stakeholders" (p. 8).

---

[44] Note that the benchmarking system was only *noted* (not adopted). As noted by Clegg (2015), this represents a lower-order form of acceptance as it does not require a positive vote.

[45] In addition, and as noted by Fukuda-Parr and Hulme (2009), the institutionalization of the MDGs went beyond the UN system – a number of donors and other development organizations set up dedicated structures, teams or projects with a focus on the MDGs.

[46] This follow-up mechanism was accompanied by the annual reporting on progress towards MDG on the part of each country and to the UN General Assembly.

*2.4.6. A note on the impact and reception of the MDGs*

While progress towards the achievement of MDGs has frequently been portrayed as uneven and at best modest (cf. Manning, 2009; and Waage et al., 2010, for an overview), the goals represented a turning point in developing thinking in that they generated an unprecedented consensus around the notion of poverty reduction, to the point of acquiring the status of a "rallying point" for the development community (Manning, 2009; see also Hulme and Scott, 2010). Some observers have noted that, while the MDGs added to an established tradition of goal-setting, they represent an element of novelty in that they achieved an unprecedented level of visibility (Manning, 2009). Such dominance of the MDGs as a global development agenda has been explained by their (unprecedented) political legitimacy secured through the MD, by their reliance on quantitative and time-bound targets, and by the articulation of an explicit and carefully arranged strategy of dissemination.

Despite their communicative success, the MDGs were not free from criticism, being challenged both on procedural and substantive grounds. In relation to the former, and as noted by Bissio (2003) and Fukuda-Parr and Hulme (2009), a number of international NGO networks were reluctant to embrace the MDGs and emphasized the undemocratic nature of their formulation process, noting that the lack of consultation during the drafting process was seen as a step back from the democratization of UN decision-making process reached during the 1990s. The very content of the MDGs was also questioned – a range of constituencies were critical of the exclusion of certain goals (especially, reproductive health), the limited ambition of others (especially Target 1A with the requirement to halve extreme poverty instead of eliminating it), and the sense of vagueness surrounding MDG 8 (the only one directly speaking to the Global North) (Reddy and Kvangraven, 2015; Saith, 2006). Some observers argued that the MDGs encouraged an undue emphasis on the meeting of basic needs and, consequently a neglect or narrowing of broad agendas set by the conferences of the 1990s, much more attentive to a human-rights approach. According to this line of reasoning, the MDGs would have "disrupted" the process of change set up by the 1990 conferences through the cherry-picking of globally-agreed priorities (Fukuda-Parr et al., 2014). In synthesis, while the UN touted the MDGs as an instance of *consensus* between the UN and the Bretton Wood institutions, the agreement was perceived as a case of *co-optation* within a part of the NGO sector. Such reservations and sense of mistrust only accrued with the appointment of Jeffrey Sachs (well-known champion of the so-called "shock therapy" as applied to post-communist states during the 1990s) as the Director of the Millennium Project.

Finally, a third strand of criticism is that concerning the monitoring strategy. The question has been discussed extensively by Fukuda-Parr (2014), who argue that the emphasis placed on simplicity and measurability, and especially the decision to only include targets with consensual indicators and robust available data, had tended to narrow the focus and the very spirit of the goals. It has also been observed that, while the simplicity and measurability attributes made the targets apposite for global mobilization and advocacy

purposes, these very attributes render them inappropriate for national planning purposes. Such issues are discussed by Jerven (2013), who has drawn attention to the perverse effects of the data demands imposed on national statistical offices by the MDGs. The author notes that most of the indicators listed in the MDGs rely on resource-intensive data-collection procedures, and that, in the context of low-income countries, the mismatch between local capacity to generate robust data and the data demands associated with MDG can lead to an inadequate use of available resources. This is true for Sub-Saharan Africa, where the authors find that:

> In some cases the monitoring demands of the MDGs have meant a windfall of economic resources for statistical offices. National accounts divisions have complained that this means that personnel from already understaffed divisions are being transferred to sections that collect data for the MDG indicators. National stakeholders, such as central banks, have said that they suspect that the quality of the important economic growth data has been decreasing […] The concern is that the limited capacity of the statistical offices is further constrained by the Millennium Development Goals agenda. (2013, p. 96)

**Wrapping up**

This chapter has outlined the emergence and consolidation of goal- and target-setting in education by reviewing a number of international initiatives relying on such devices in order to symbolize a sense of shared purpose and galvanize collective action. In so doing, the chapter has highlighted the simultaneously institutionalized and open-ended character of such practices. On the one hand, most of these exercises rely to some extent on a "toolkit" or playbook of codified procedures and solutions that offer some guidance when it comes to drafting procedures, consultation strategies, consensus-making mechanisms, or the very selection of participants. On the other hand, this "toolkit" does not operate as a blueprint in the sense that each one of the reviewed conferences, declarations and goal-setting exercises has put forward a number of procedural innovations and has effectively "broken ground" in one way or another.

The chapter has also identified a number of recurrent tensions and trade-offs in target-setting as a global governance practice. First, a tension between the use of broad sets of goals formulated in a relatively ambiguous way and consequently capable of operating as rallying points for the international community, and the use of more limited sets of actionable goals more likely to set clear priorities and induce change. Secondly, the design of follow-up or implementation mechanisms also appear to pose a number of challenges. Although such mechanisms are increasingly recognized as necessary to prevent goals and declarations from becoming mere wish lists, they generally demand the creation of new mandates (or the clarification of the existing ones), thus rendering evident the existence of power asymmetries and organizational conflict, and the limiting impact of institutional inertias. Finally, the degree of ambition of the goals and targets, as well the level of clarity in the attribution of responsibilities, represent recurrent issues of contention for which creative solutions are needed.

# Theoretical antecedents and conceptual considerations

## Introduction

While goal- and target-setting are well-established practices within the development realm and within the UN system in particular, they remain the subject of a rather modest body of literature. To be sure, and as noted in Chapter 2, there are a number of fine-grained, richly detailed accounts of the negotiation of SDGs, MDGs and other goal sets. Yet, insightful as they are, such accounts leave a number of issues open – particularly regarding the nature of global goals and targets as an instrument or technology for global governance. To put it differently, existing literature lacks a theoretical articulation of the nature and specific properties of global goals as tools of soft power. Within much of the available scholarship, this is a question that remains implicit rather than explicitly theorized.

Goal-setting thus seems to be an elusive, conceptually challenging object of study. This is largely the result of the one-off and highly idiosyncratic character of exercises of collective commitment, such as the one goal-setting represents. Hence, global goal-setting practices resist straightforward generalization. Matters are further complicated by the fact that, as noted above, SDGs represent a form of discontinuity regarding earlier forms of goal-setting – given the unprecedented centrality of quantitative targets that characterize the new agenda. Numeric indicators constitute the bedrock of the new agenda, rather than a mere add-on to the declarations of shared aspiration typically associated with international summitry. Therefore, SDGs are likely to trigger a series of specific mechanisms that were less prominent in other global agendas.

Since relatively few scholars have focused upon global goal-setting as a specific object for theoretical inquiry, this chapter engages with two related areas of inquiry – one focusing on processes of international norm-building and the other addressing the making of global indicators. These two areas of inquiry correspond to two distinct (albeit interrelated) properties of SDGs frequently highlighted by specialized literature – namely their potential to trigger normative shifts (thus working as an instance of global norm setting), and their constitutive effects (that is, their capacity to redefine reality through a series of measurement choices). Together, both literatures provide the basis for theoretical engagement with the concept of global goals, and for articulation of SDGs as an object of research.

The chapter explores a third relevant theme for goal-setting analysis – the question of organizational interdependence. This addresses the paucity of theoretically driven accounts discussing how inter-organizational dynamics shape (and are shaped by) the practice of global goal-setting. This theoretical problem is of particular importance given

the research's interest in interorganizational negotiations leading to adoption of SDG4, as well as on different forms of organizational interdependence brought about by the new monitoring needs. Because SDG4 was born out of multiorganizational dynamics, and its monitoring and implementation bring together different organizations, the chapter identifies those theoretical constructs most appropriate to capture and make sense of inter-organizational interaction.

Thus, the main objective of this chapter is to review the theoretical antecedents of the three main themes around which this dissertation revolves, namely norm-making processes; the development of global indicators; and organizational interdependence. For each one of these, the chapter identifies and elaborates those strands of literature more appropriate for refining and bringing into sharper focus the questions addressed by the dissertation.

## 3.1. Understanding norm-making processes

### 3.1.1. Approaching goal-setting as an instance of norm-building

This dissertation approaches global goals and targets as instances of global norm-building – that is, as one of the vehicles or products through which global norms are manufactured and crystallize. This approach is coherent with and follows on from the work on MDGs conducted by development scholars Sakiko Fukuda-Parr and David Hulme – who feature among the few theorizers of global goal-setting practices. In the seminal paper *International Norm Dynamics and the "End of Poverty": Understanding the Millennium Development Goals* (Fukuda-Parr and Hulme, 2011), the authors examine MDGs through the lens of international norm dynamics as developed by Finnemore and Sikkink (1998) (see section 3.1.2 below) while proposing some conceptual adjustments to make sense of the specificities of this particular set of goals. The authors thus argue that MDGs should be understood as constitutive of a *supernorm*, that is, as "carefully structured sets of interrelated norms that pursue a grand prescriptive goal" (2011, p. 31). This is so that, while each of the eight MDGs functions as a distinct, separate norm, they are integrated in an umbrella declaration, and they list and (more importantly) exhibit a certain coherence in terms of ends and means, and are held together by a common thread – namely, the need to eradicate extreme poverty, thus encapsulating a wide range of commitments and agreements reached through a decade of global summits and conferences.

More recently, Fukuda-Parr and McNeill (2019) have refined and added nuance to this early conceptualization of global goals as instances of norm-building. In their account of negotiation of the new development agenda advanced by SDGs, the authors argue that global goals need to be understood as *instruments* or *vehicles* for internationally agreed norms – rather than norms themselves. The authors thus note that global goals "are a particular institutional arrangement for norm creation (elaborating, negotiating, achieving

consensus), institutionalization (communicating and diffusing the norm), and implementation (evaluating performance and creating incentives)" (p. 6). As such, they contribute to shaping the priorities of both states and other development stakeholders, ultimately shaping the normative evolution of development thinking.

A *solo* paper by Fukuda-Parr (2011) also provides some useful insights regarding the specific properties of global goals as a form of norm-building. In particular, the author draws attention to the specificities of MDGs as essentially *aspirational* norms, devoid of a clear implementation strategy and, more importantly, dissociated from a clear set of policy rules. Fukuda-Parr (2011) reflects thus on the limited political traction of MDGs among the donor community, drawing attention to the fact that normative consensus can coexist with a lack of policy change – and that this is likely to occur when there is no agreement (or reflection) on the causal forces behind the targeted issue. A similar point is made by Finnemore (2007), who notes that MDGs constitute a particular variety of commitment, likely to be perceived as "UN-sponsored platitudes" – in that they gather support at the rhetorical level but create weak incentives for many governments (for whom the political payoff from implementation of MDGs is, at best, unclear), as well as for many IOs (for whom the MDGs constitute "add-on" objectives rather than part of their core mission) (Finnemore, 2007). These observations suggest the need to examine MDGs (and, by extension, SDGs) as an extreme example of a *thin norm* (cf. Wiener, 2003) – i.e., a norm which is able to accommodate multiple interpretations but with a problematic or unclear translation into policy rules.

However, the limited political traction associated with global goals should not be equated with inconsequentiality. Fukuda-Parr and Yamin's (2013) paper on the normative trajectory of MDGs, along with Fukuda-Parr and McNeill's recent (2019) account of the negotiation of SDGs, observes that the ultimate impact of MDGs and SDGs is not only a consequence of their impact on the policy priorities of key stakeholders but also of their capacity to redefine and reinterpret understanding of key themes within the development realm. Similarly, in her comment on MDGs, Finnemore draws attention to the fact that "symbolic politics is not empty politics" (2007, p. 105), that is, that the symbolic and rhetorical commitments advanced by such goals *do matter* – for instance, as they are likely to become focal points around which NGOs or the general public mobilize, and to condition the ascendance of certain issues within organizational policy agendas. Overall, such observations suggest that, regardless of their ultimate political impact, global goals represent meaningful and consequential instances of international norm-building and need to be articulated as such.

### 3.1.2. Conceptual considerations on the notion of norms

Despite the ubiquity of the term within IR scholarship and policy studies, there is no univocal understanding of the notion of norm. The rationalist vs. constructivist (or cognitivist) divide features among the most relevant breaches (Checkel, 1997;

Wunderlich, 2013). Rationalist approaches emphasize the constraining effects of norms and understand norm adherence as a reflection of actors' cost-benefit calculations. Conversely, perspectives falling within the cognitivist paradigm (including constructivist ones) emphasize instead the key role of norms in the formation of agents' identities and preferences, and approach norm compliance as driven by an inclination to comply with the obligations or duties associated with a role, identity or membership in a community – thus invoking the so-called logic of appropriateness (cf. March and Olsen, 1989). Realist and neorealist theories, in turn, understand norms as resulting from an attempt to justify coercive dynamics and dominance exerted by a group of powerful states. Norms are thus conceptualized as an epiphenomenon – that is, a mere by-product, devoid of explanatory power (Checkel, 1997; Florini, 1996).

For the purposes of the research, this dissertation relies on a constructivist understanding of norms. The conceptual precisions advanced here are therefore informed by this specific perspective. Thus, from a constructivist perspective, and according to Katzenstein's oft-quoted definition, norms are conceptualized as "collective expectations for the proper behavior of actors with a given identity" (1996, p. 5). Another widely accepted definition is the one advanced by Finnemore and Sikkink, which approaches norms as the "appropriate behavior for actors with a given identity" (1998, p. 891).

Broadly speaking, and as these definitions suggest, the constructivist perspective on norms is characterized by an emphasis on a number of distinct qualities, including (*a*) the collective and intersubjective dimension of norms – that is, the fact that norms are collectively shared and constructed through social interaction; (*b*) their constitutive dimension[47] – that is, on the fact that norms fix meanings, create categories of actions and actors, and contribute to shaping actors' identities, interests and beliefs; and (*c*) their evaluative dimension – that is, their prescriptive character or the so-called "quality of oughtness", in other words, their dependence on some form of (collective) moral assessment (Finnemore and Sikkink, 1998; Wendt, 1999; Wiener, 2003).

As a result of the above, for a norm to exist, it must be recognized to some degree by relevant stakeholders. This collective dimension is, in turn, the product of its institutionalization – that is, "the perceived legitimacy of the norm as embodied in law,

---

[47] As noted in this section's opening paragraph, the constitutive dimension is frequently drawn in opposition to the constraining dimension emphasized by neorealist theories and by rational-choice perspectives – that is, the idea that international norms constrain behavior or domestic decision-makers by altering their calculations about the costs and benefits of a given course of action (Florini, 1996). Such propositions have, however, been contested by Checkel (1997), who observes that constitutive and constraining effects are not mutually exclusive, and that constructivist scholarship would benefit from greater attentiveness to the latter. The varying degrees of attention given to the question of constitutive effects have indeed contributed to the emergence of a breach between so-called conventional and critical constructivists (Wiener, 2009). This dissertation remains agnostic on such issues as they do not have direct implications for the conceptualization of the object of study, which is focused on the *emergence* and *negotiation* of norms rather than on their effect on domestic constituencies.

institutions or discourse, even if all relevant actors do not follow it" (Onuf, 1997 in Bierman, 2013). However, normative compromises such as those advanced by norms do not necessarily translate into a set of standardized rules of behavior – on the contrary, most international norms are characterized by a considerable degree of ambiguity and flexibility. This is so as the legitimacy enjoyed by a norm and the identification it generates among its addressees are generally proportional to the scope for interpretation the norm allows for (Wiener, 2003). As noted by Bierman (2013), this is likely to be the case with those compromises that obscure differences and mask trade-offs, rather than exposing or confronting different views. Such dynamics have a direct impact on not only country compliance but also the action of different multilateral organizations – which are likely to leave their core missions unaltered and adopt the new discourse only at a rhetorical level, using them as "window-dressing" (Tallberg, Lundgren, Sommerer and Squatrito, 2018). Similarly, Finnemore and Hollis (2016) note that formal adherence frequently coexists with more or less explicit contestation, including so-called insincere conformity (when relevant stakeholders "pay lip service" to the norm but do not modify their behavior) or incompletely theorized agreements (when relevant stakeholders agree on what constitutes appropriate behavior but not on the reasons why this behavior is desirable).

Empirical observation of norms, therefore, cannot be (solely) based on the identification of behavioral regularities among those targeted by them. Norms can, however, be operationalized since, as noted by Bernstein (2013), "[N]orms leave behavioral traces in the form of treaty commitments, action programs, outcomes documents, practices, policies, and so on" (Bernstein, 2013, p. 128). Such questions have also been addressed by Halliday, Block-Lieb and Carruthers (2009), who have proposed the notion of script as "tangible evidence of global norm-making". The authors thus observe that when IOs engage in creation of global norms, they articulate them in different textual products – including those resulting from multilateral conventions, legislative texts and standards etc. International norms crystallize and are thus codified in a variety of formats whose formal and rhetorical properties can be analyzed. A similar point is made by Tallberg, Lundgren, Sommerer and Squatrito (2018), who note that norm acceptance (but also norm contestation) can be operationalized and empirically observed, again, through those documents in which normative principles are explicitly stated, as well as in the justifications and accounts produced by deviating actors.

### 3.1.3. The problem of norm emergence

Along with the notion of identity, norms rapidly became one of the cornerstones of the constructivist research program consolidated during the late 1980s and early 1990s. By bringing to the forefront the key role of ideational forces and intersubjective knowledge, as well as the socially constructed nature of reality, research on norms came to be seen as a means of exposing the shortcomings of the neorealist and neoliberal schools then dominating the IR field (Björkdahl, 2002; Hoffmann, 2010). Thus, a range of works now

seen as foundational for early constructivist theory directly engaged with the notion of norm, exploring its theoretical boundaries and using it as a key analytical device to explain a variety of phenomena (see, for instance, the pioneering works of Finnemore, 1996; Finnemore and Sikkink, 1998; Katzenstein, 1996; Keck and Sikkink, 1998; Klotz 1995; and Risse, Ropp and Sikkink, 1999).

This initial push on norm-oriented research tended to focus on norm compliance or norm conformity – thus conceptualizing norms as variables with independent explanatory power, rather than phenomena to be explained (Björkdahl, 2002), and approaching them as stable, fixed and static entities (Wunderlich, 2013). The emphasis on norm-induced behavioral change was driven by the need to prove that norms mattered and constitute valid alternatives to the rationalist and materialist variables emphasized by more established theories (Checkel, 1997; Hoffmann, 2010). This first wave of literature also addressed the question of norm spread (and the associated problem of socialization dynamics) but failed to give a comprehensive and nuanced account of processes of norm change, as well as of norm emergence and/or decay (Sending, 2002; Wunderlich, 2013).

To be sure, this first wave of norm-oriented research put forward some models aimed at capturing the "life cycle" of norms, thus identifying norm emergence as a distinct moment worthy of scholarly attention. This is particularly the case with Finnemore and Sikkink's (1998) three-step process, which draws a distinction between norm emergence, norm cascade and norm internalization. However, empirical research on norm emergence has been limited and generally restricted to instances of norm entrepreneurship initiated by a few "external agents", that is, non-state and non-governmental actors – including transnational advocacy networks (Keck and Sikkink, 1998) and experts articulating in epistemic communities (Haas, 1992).

Thus, despite some initial progress, the question of norm emergence tended to remain comparatively underdeveloped for a considerable period. Finnemore and Sikkink have noted critically that scholarship on the influence of global norms "sometimes overlooks the fact that international norms have to come from somewhere" (2001, p. 397). Similarly, and tacking stock of the theoretical advances facilitated by the constructivist turn, Wiener observed in 2003 that research on norms has been driven by two main sets of questions, which have received unequal attention. The author noted that the main line of inquiry revolved around the question of norm impact, engendering extensive discussion of the varying degrees and rationales of state compliance. Conversely, the very emergence of norms, that is, the process through which these sets of shared principles come into being, has remained less systematically examined (Wiener, 2003).

The situation started to change around the turn of the century, with a number of works turning the spotlight on the *dynamic* character of norms (Sandholtz, 2008; Wunderlich, 2013). These new accounts started to challenge the (then prevailing) lineal understanding of norms, paying greater attention to their evolving nature, as well as to episodes of norm contestation and the interactive nature of processes of norm construction (Hoffmann,

2010; Wiener, 2007). Overall, such shifts helped to add nuance to the processes by which certain ideas eventually acquire a normative status. Importantly, this second generation of constructivist-oriented research entailed a diversification of the concept of norm entrepreneurs and, more generally, a richer understanding of the nature of the agents driving norm emergence or change. The initial emphasis on the role of non-state actors (including NGOs, social movements and epistemic communities) was thus complemented by greater attention given to state entrepreneurship (Björkdahl, 2002; Wunderlich, 2013), as well as to supranational entrepreneurship – that is, the role of international bureaucracies (Tallberg et al., 2018). As noted by Kentikelenis and Seabrooke (2017), global norm-making is today understood as an iterative process, drawing together different combinations of states, IOs (including intergovernmental as well as international non-governmental organizations) and professional groups, who interact and vie for the advancement of their normative preferences (see also Halliday and Carruthers, 2007; and Halliday et al. 2009).

### 3.1.4. Understanding the role of international organizations

Although the key role of IOs as "teachers of norms" had been recognized early on by Finnemore (1993), empirical and theoretical work on the role of international bureaucracies in normative processes did not take shape until the mid-2000s. A key driver of change was growing scholarly interest in the role IOs as autonomous organizations, spearheaded by Barnett and Finnemore (1999, 2004),[48] which allowed for greater theoretical elaboration of the role of IOs in the spread of global norms. In their seminal essay *The Politics, Power, and Pathologies of International Organizations* (1999), Barnett and Finnemore identified the diffusion of global norms as one of the key mechanisms through which IOs exercise power and position themselves as sites of authority – along with classification and the fixing of meanings.[49] The authors noted:

> Having established rules and norms, IOs are eager to spread the benefits of their expertise and often act as conveyor belts for the transmission of norms and models of "good" political behavior […] Armed with a notion of progress, an idea of how to create the better life, and some understanding of the conversion process, many IO elites have as their stated purpose a desire to shape state practices by establishing, articulating, and transmitting norms that define what constitutes acceptable and legitimate state behavior. (p. 712–713)

Barnett and Finnemore's work thus portrayed IOs as both transmitters but also as creators of norms – although most of the empirical discussion revolved around the role of these

---

[48] In this work, the authors argued in favor of examining IOs as (autonomous and active) bureaucratic organizations, drawing attention to the theoretical significance of the notion of *organizational culture* – defined as the "rules, rituals, and beliefs that are embedded in the organization (and its subunits)" (Barnett and Finnemore, 1999, p. 719).

[49] Note that the authors returned to this classification in the 2004 book *Rules for the World* – see Chapter 2 ("International Organizations as Bureaucracies").

organizations in norm diffusion. The agency exerted by IOs as shapers of their own normative environment was also recognized in Barnett and Coleman's discussion of IOs as strategic actors (2005). The authors observed that IOs can turn to strategies of manipulation and strategic social construction as a means to ensure their own survival and preserve their mandates and autonomy. IOs are thus likely to engage in transformation of the normative environment in which they operate in order to secure some degree of consistence with the organizations' goals or principles. The authors concluded that "Although traditional approaches to IOs have tended to assume that the environment dominates IOs, our approach suggests that IOs may choose to accept, resist or actively change environmental pressures" (p. 602).

However, even if these insights were instrumental in bringing to light the normative labor performed by IOs, theoretical and empirical progress on such questions was not immediate or straightforward. Despite the growing consensus among constructivist quarters on the fact that IOs do not merely espouse norms but actively *produce* them, scholarship examining how such a process unfolds remained limited until the mid-2000s. In an overview of early works on IO engagement with norms, Park (2006) observed that the sources of IOs' preferences or normative approaches were rarely addressed in much depth. Rather, IOs were portrayed as norm diffusers or norm transmitters, while the ultimate origin of their normative orientation, as well as their potential role as norm-makers, remained obscured or unaddressed. Park thus advocated for a new line of inquiry, noting that "A more thorough account of how such norms emerge to dominate an IO and reconstitute its identity therefore needs to be discerned in order to understand why an IO's identity changes and what relation, if any, this has to the role of IOs as norm diffusers" (2006, p. 353). The author thus argued that there was insufficient understanding regarding both the emergence of norms within IOs and the interaction of IOs with external actors during norm-building processes.

Similar concerns were voiced by Chwieroth (2008), who observed that the emphasis of much constructivist literature on the significance of IOs' organizational culture had come at the expense of an understanding of the proactive role of IO staff in processes of normative change, as well as of the dynamics of internal contestation. The author noted that much of IO literature tended to focus on "external" norm entrepreneurs while downplaying the role of norm entrepreneurs operating *from within*. As a result, constructivist literature would have tended to emphasize staff resistance to normative change – largely in order to substantiate the theoretical proposition that organizational culture and routines do matter.[50] In consequence, staff behavior would have entered the picture as uniform, static and essentially reactive.

---

[50] Well-known resistance episodes include, for instance, Momani's (2005) study of the IMF staff's response to attempts made by the Managing Director and Executive Board to reform the process of design and negotiation of conditionality in order to favor country ownership. The author shows that top-down calls for change were resisted by IMF personnel, eventually having little impact at the operational level and being instrumentalized to secure greater staff discretion. Similar dynamics are described in Weaver and Leiteritz's (2005) study on the World Bank's Strategic

Such trends have been at least partially reversed over the last decade – particularly as scholarship on IOs has become increasingly concerned with endogenous drivers of organizational policy change (Vetterlein and Moschella, 2014). There is thus growing evidence that IO staff actively engage in different forms of norm entrepreneurship, shaping the institutional agenda and advancing normative shifts. This is, for instance, the case of Leiteritz and Moschella's (2010) study on the trajectory of the capital account liberalization policy norm. The authors note that consensus around the idea initially developed *inside* the IMF during the 1980s as a result of the advocacy efforts of staff members. Thus, the authors explicitly reject the notion that the norm was imposed on the organization by member states acting in accordance with the private sector. Similarly, in a study of the so-called gender and development[51] (GAD) policy norm, Weaver (2010) finds evidence that the institutionalization of this perspective within the World Bank was largely the result of the lobbying efforts of a number of norm entrepreneurs operating within the organization, who relied on a strategic framing tactic to demonstrate the organizational "fit" of the GAD approach and made sure it resonated within the prevailing organizational culture.

One of the most frequent themes raised by studies on processes of (internal) normative change is the key role of expert authority. Specialists and experts are often seen as key sources of change within IOs, both in terms of operational policies and organizational priorities and objectives. This is, for instance, the case with Broad's (2007) research on the World Bank's DEC,[52] in which the author argues that the World Bank research unit played a crucial role in articulation, maintenance and projection of the neoliberal economic paradigm that came to dominate development thinking and practice during the 1980s and 1990s. This is consistent and resonates with other works focusing on processes of policy change within IOs, particularly with those emphasizing the role of ideas as drivers of such transformation. An example of this is Béland and Orenstein's (2013) discussion of the role of individual policy experts operating within IOs. Focusing on evolution of the internal discourse of the World Bank on pension privatization, Béland

Compact reform initiative, launched by President James Wolfensohn. Studies of so-called *organized hypocrisy* have also tended to emphasize IOs' *reactive* behavior and episodes of inertia and resistance. Such studies examine how IOs and other organizations frequently respond to conflicting pressures in their external environment by pulling in different and sometimes opposing directions. One of the consequences of such processes is a decoupling between talk (recently espoused mandates or goals) and action (oriented by preexisting routines, established procedures and entrenched staff expectations etc.) (see, for instance, Lipson, 2007: and Weaver, 2008). While these studies do not preclude the possibility of internally driven normative change (and in fact grapple with episodes of internal policy advocacy), such accounts tend to emphasize the enduring quality of organizational culture.

[51] GAD approaches emphasize the need to mainstream gender in development programs across all policy sectors and intervention levels.

[52] The Development Economics Vice-Presidency (DEC) is conventionally considered to be the World Bank's main research unit, headed by the Chief Economist and entrusted with intellectual leadership of the institution.

and Orenstein (2013) contend that, at least in those organizations deriving their authority from expert knowledge, strategically located individual experts affiliated to the organization might operate as policy entrepreneurs and contribute to shape and transform the policy priorities or ideas espoused by IOs. The key role of experts has also been captured by the work of Broome and Seabrooke (2012) on so-called analytic institutions, defined as "specialist units, departments, committees, adjudicatory bodies and others housed by or linked to IOs that develop the cognitive framework for understanding and solving policy problems" (p. 3). The authors argue that such expert units play a crucial agenda-setting role by diagnosing (and proposing solutions to) global and domestic problems.

Another strand of literature has focused on the role of organizational leaders as key instigators of institutional and normative change in global policy spaces as well as within IOs. Although this has long been an undertheorized question in the study of IOs, some scholars have recently started delving into the influence wielded by high-ranking officials in a more systematic way (cf. Kille, 2013, for an overview of the role of Secretaries-General within the UN system; and Hall and Woods, 2018, for a comprehensive state of the art). Insightful studies include Schroeder's (2014) analysis of two different executive heads – President R. McNamara of the World Bank and the UN Secretary-General D. Hammarskjold. The author argues that the room for maneuver enjoyed by top officeholders is mediated by both organizational and environmental conditions, and identifies two ideal-type mobilization strategies available to top leaders – a follower-oriented strategy (targeted on those agents inclined to second the proposed changes) and an opposition-oriented one (oriented towards preventing the emergence of "blocking coalitions").

The notion of organizational subcultures adds yet another layer of complexity to the norm-emergence debate, as it has drawn attention to the plural character of international bureaucracies as sites of normative change. The significance of professional divergences and internal dissent had indeed been anticipated by Barnett and Finnemore's (1999, 2004) elaboration of the notion of *cultural contestation* – portrayed as a form of pathological behavior resulting from the division of labor inherent in bureaucratic organizations and the subsequent emergence of "local cultures" exhibiting competing views. The authors observed:

> Organizational culture is an accomplishment rather than a given. Organizational control within a putative hierarchy is always incomplete, creating pockets of autonomy and political battles within the bureaucracy […] Different segments of the organization may develop ways of making sense of the world, may experience different local environments and receive different stimuli, and may be populated by different mixes of professions or shaped by different historical experiences. (2004, p. 40–41)

More recently, Nelson and Weaver (2016) have argued that subcultures and countercultures are likely to emerge "within large organizations with broad or multiple

mandates, where there are likely to be several different staff specializations or professions, as well as high staff turnover" (p. 925).

More empirically driven accounts of such subcultures include the work of Chwieroth (2008, 2010), who has argued that organizational subcultures are likely to emerge in those organizations drawing from a wide range of professions, and that such subcultures tend to engender "battles of ideas" mirroring different professional beliefs. The author finds that the professional diversity that has characterized the World Bank since its inception has engendered fiery debates among the staff, translating into interdepartmental conflict and different forms of bureaucratic infighting between lawyers, bankers and economists. Similarly, Bebbington, Guggenheim, Olsen and Woolcock (2004) also observe that "[*D*]ifferent groups in the [*World*] Bank have different languages" (p. 44). The authors draw thus a distinction between an first group of non-operational staff working on policy analysis and research (including DEC staff, country economists and country managers), who would privilege an "econometric grammar" and quantitative procedures oriented towards generalization, and a second group composed mainly of operational staff more concerned with data usability, and inclined to work with qualitative data and context-specific models informed by political economy.

Beyond these few exceptions, however, much of the empirical scholarship on the impact of professional cultures has tended to emphasize ideological consistency and disciplinary uniformity[53] – rather than exploring internal friction or the role of non-dominant professional groups. An additional difficulty lies in the fact that, as the examples above suggest, most of the empirical literature on professional cultures has focused on Bretton Woods institutions. Overall, the existence of intra-organizational subcultures remains more of a theoretical possibility than the object of systematic investigation – and the impact of such organizational subcultures in relation to norm-making procedures thus remains largely undertheorized.

*3.1.5. Relevant analytic parameters*

As per the previous section, most of the studies discussing normative change *within* IOs adopt a thick-description approach, focusing on issue-specific and organization-specific dynamics. Echoing the idiosyncratic nature of existing accounts (and the fact that they constitute a rather fragmented corpus of research), there is a dearth of literature identifying organization-invariant dimensions of analysis. Park and Vetterlein's (2010) analytical parameters for analysis of norm change within Bretton Woods institutions are

---

[53] This is, for instance, the case with Rao and Woolcock's (2007) discussion of the disciplinary near-monopoly held by the economics profession within the World Bank, or Babb's (2003) observations of the influence wielded by economists as a key driver of the IMF's shift towards the neoliberal model in the 1980s. Similarly, Nelson (2014) finds evidence of a certain degree of professional homogeneity within the IMF – with a significant share of staff having been trained in economics in Anglo-American institutions.

an exception to this. Drawing together the results of a range of empirical studies examining processes of norm emergency and change within the World Bank and the IMF, Park and Vetterlein (2010) identify two main axes along which the emergence and consolidation of norms within IOs can be examined, namely:

> *(a)* Change from inside IOs vs. change from outside. At one end of the spectrum, we find norm advocacy on the part of external actors (including non-state actors such as NGOS but also other IOs), whereas at the other end, we find norm advocacy driven by IO staff, that is, from international bureaucracies. The authors place the role of member states in an intermediate position – given that they are in a privileged position to affect the direction of the organization but only through delegated powers.

> *(b)* Top-down vs. bottom-up sites for policy-norm emergence. The continuum ranges from those cases in which norms are advanced from member states to management and from management to staff, to those cases in which new norms come from staff and/or from non-state actors and only then ascend to management positions.

A second and complementary contribution to the analytic framework sketched by Park (2005) is the so-called *syncretic approach* to the emergence and institutionalization of scripts within IOs, advanced by Kentikelenis and Seabrooke (2017). These authors draw attention to the need to unpack dynamics of scriptwriting (that is, codification of norms into prescriptive behavioral templates) within IOs, arguing that such dynamics remain comparatively undertheorized. The authors thus propose an analytical model oriented towards capturing different forms of interaction between organizational staff and the representatives of member states, and add a layer of complexity by incorporating into the model the specific power hierarchy exhibited by each one of these groups. In the case of country representatives operating through the board of directors (or an analogous structure), power asymmetries among different delegates and states are a function of a combination of factors, including voting rights as mediated by voting procedures, contributions and ascendancy over budgetary decisions, coalition-building dynamics and informal conventions (*i.a.*). In the case of organizational staff, power and authority are distributed according to (organization-specific) hierarchical structures but are also affected by other considerations, including scientific expertise and prestige, with positions generally clustering into three main groups: senior managers, policy staff and young professionals.

Finally, the authors contend that the interaction between member states and staff is mediated by two main factors, namely (*a*) the degree of attention and oversight over policy activity on the part of member states – ranging from close monitoring to indifference or laxness; and (*b*) the nature of policy activity conducted by staff – ranging from what the authors term "consistent policy activity" (practices in line with already-consolidated scripts, or in accordance with preferences expressed by the member states)

to "mixed or inconsistent policy activity" (including attempts to introduce and institutionalize new norms). On this basis, Kentikelenis and Seabrooke (2017) anticipate four possible outcomes of intra-organizational interactions. These include (*1*) consensual dynamics (when the preferences of different tiers coincide); (*2*) contentious scriptwriting (when staff-driven policy changes are met with resistance or blockage on the part of powerful states, or when there is some form of infighting among the staff or among country delegates); (*3*) script stability (driven by a high degree of oversight on the part of the board); and (*4*) the prevalence of ad hoc decisions (resulting in a combination of indifference on the part of country representatives and the absence of a unitary direction among the staff).

## 3.2. The making of global indicators

### 3.2.1. Understanding SDGs as a quantification exercise

As discussed in the introduction to this chapter, the transformative capacity associated with global goals appears to be magnified by the growing reliance of such goals on quantitative targets. In the UN system, this shift has been initiated during the MDG era and consolidated with approval of SDGs. As noted by Fukuda-Parr and Yamin (2013), the incorporation of numeric indicators into MDGs is at the root of their symbolic and discursive impact, as well as of their awareness-raising potential. Measurement efforts are thus deemed central to explaining the unexpected (and sometimes undesired) knowledge effects of MDGs – that is, their impact on conceptualization of development issues. Similarly, Clegg (2015) argues that MDGs need to be understood as one of the most prominent (global) benchmarking exercises in the history of development. The introduction of indicators would thus represent a qualitative change in the practice of global goal-setting, in that it redefined the distribution of responsibilities and succeeded in altering the balance of power.

More recently, the unprecedented centrality of quantitative targets, characteristic of SDGs, has been highlighted as indicative of a methodological shift which has vast implications for knowledge and political dynamics shaping the development realm. As a consequence of the growing significance of measurement tools in giving content and meaning to SDGs, Fukuda-Parr and McNeill have noted that "[*T*]he real locus of power in setting international agendas has shifted to the selection of indicators. The exercise of power takes place through multiple steps in the process of setting the goals and measuring them" (2019, p. 14).

A similar point has been made by the quantification theorist Sally Merry (2019), who has observed that SDGs' reliance on quantitative indicators (as a key technology with which to foster and monitor progress) entails a fundamental transformation of the existing understanding of development. The author adopts a critical stance, noting that the central role assigned to indicators is particularly problematic given the many constraints of the

measurement infrastructure currently in place (i.e., the expertise and resources available with which to quantify a phenomenon of interest). Merry (2019) concludes that these limitations (including limited funding, data-collecting and methodological difficulties, and the imperfect nature of available proxies) are likely to erode the more aspirational dimension of SDGs.

Such observations suggest that an understanding of quantification practices is necessary to make sense of the ultimate impact of those global goals characterized by a heavy reliance on quantitative indicators – as is the case with SDGs. In other words, measurement choices need to be brought to the forefront if we are to understand the normative impact of global goal-setting.

### 3.2.2. The proliferation of global indicators

Indicators play an increasingly prominent role in global governance. Indeed, the rapidly growing availability of global indicators has been compared to an explosion or avalanche (Bhuta, Malito and Umbach, 2018) whose origins can be traced back to the end of the Cold War – although production of such metrics has significantly accelerated over the last two decades (Cooley, 2015; Kelley and Simmons, 2014). Such proliferation of global indicators is discussed by Merry (2016) as being one of the many manifestations of the emergence and consolidation of an *indicator culture*[54] – in turn, a variation of the audit culture in the sense advanced by Shore and Wright (2015). This indicator culture is defined as a series of practices and assumptions about knowledge production, characterized by their "trust in technical rationality, in the legibility of the social world through measurement and statistics, and in the capacity of numbers to render different social worlds commensurable" (Merry, 2016, p. 9). As noted by the author, the emergence of the indicator culture is consistent with consolidation of so-called evidence-based governance – that is, those regulatory strategies that put a premium on data-based approaches to decision-making, and that privilege the use of standards, evaluation, monitoring and self-assessment practices over the imposition of sanctions and/or top-down command. Arguing along the same lines, Cooley (2015) observes that the appeal of comparative data cannot be dissociated from the fascination with performance evaluation techniques promoted by the neoliberal turn and the dissemination of new public management theories. Additionally, the author draws attention to two additional drivers explaining growth in supply and demand for *global* indicators – specifically, a number of recent changes in information technology that have greatly facilitated information compilation and data analysis processes; and the proliferation and strengthening of global governance networks and international regulatory bodies able to exert significant influence over a range of transnational issues (Cooley, 2015).

---

[54] Merry (2016) cautions against understanding indicator culture as a totalizing or holistic phenomenon, and she explicitly argues that, as with audit culture, indicator culture is integrated in the repertoire available to those actors seeking to influence decision-making procedures.

As the last point suggests, the explosion of global indicators can be understood as a precondition for the articulation of a transnational architecture of governance (Bhuta et al., 2018). This is a question of particular relevance given the purposes of this inquiry. Thus, as different scholars have argued, the centrality of global indicators for the exercise of transnational governance can be understood as analogous to the key role of statistical knowledge during the rise of the modern state – a process extensively discussed by early theorists of the social significance of quantification.

This is most notably the case with Porter (1995), who argued that the increasingly prominent role of quantitative expertise in processes of public decision-making owes much to a certain idealization of *objectivity* and its association with political democracy – for it fulfills the democratic demand for impartiality and fairness, and replaces the exercise of personal judgment and discretion typically associated with elite decision-making. Similarly, Desrosières (1998) contended that development and evolution of statistical tools and procedures were central in the articulation of a public sphere as they created a shared language to describe social and economic phenomena, necessary for public debate. Statistical reasoning thus brought into existence a series of objects otherwise hard to define or demarcate – which, in turn, worked as reference points, common to a wide range of actors.

Mirroring such processes, global indicators would thus operate as "an emerging technology in the practice of global governance" (Davis, Kingsbury and Merry, 2012, p. 6). Global indicators are thus seen as a fundamental component of global governance as they provide new forms of knowledge, necessary for the exercise of power across and beyond the boundaries of the national scale (Cooley, 2015; Davis et al., 2012). Indicators are indeed more likely to be mobilized as technologies of governance in those settings characterized by fuzzy lines of authority, ambiguous jurisdiction and a reliance on soft rather than hard law – and governed through guidelines rather than rules (including development aid, human rights and global education systems) (Merry, 2016). The constitutive role of worldwide data as a sort of basic infrastructure necessary for transnational governance is synthesized by Bhuta, Malito and Umbach (2018), who argue:

> [*J*]ust as government at the scale of the nation-state requires the production of knowledge that facilitates knowing and acting at a national territorial scale, governance at the global level demands forms of knowledge that are sufficiently stripped of national and local contexts to facilitate comparison, judgement, and action across national and regional scales. Quantification in the form of indicators facilitates the creation of such global-scale knowledge. (p. 12)

Such accounts emphasize the role of indicators as a *precondition* for articulation of a transnational architecture of governance. Other authors, however, note that the reverse argument applies too – that is, that the explosion of indicators results from a process of densification of transnational governance networks. As noted by Broome and Quirk (2015a), recent changes, including the rapid increase of NGOs in a variety of areas, have

resulted in expansion of the number of "prospective benchmarkers" (p. 820). In a context of market saturation, where a growing number of actors need to compete to secure the salience of their political causes and exert a meaningful influence over agenda-setting dynamics, the production or sponsoring of benchmarks is perceived as a privileged means of securing a position of authority and enhancing organizational credibility. As a result of these dynamics, the authors note that benchmarking can be understood as an instance of market competition for resources but also for attention, credibility and prestige.


### 3.2.3. Indicators in transnational governance: conceptual remarks

Within the literature discussing quantification practices in relation to transnational governance, the concept of an *indicator* is often used with an emphasis on its *comparative* and *evaluative* dimension – to the point that indicators, benchmarks and rankings are frequently used in an interchangeable way or even treated as synonyms. However, each one of these concepts has specific properties.

The notion of an indicator qualifies as the broadest of the concepts listed above. While there is limited consensus on its conceptual boundaries, a commonly accepted definition of the term is the one advanced by Gallopin (1996), who defined indicators as "variables that summarize or otherwise simplify relevant information, make visible or perceptible phenomena of interest, and quantify, measure, and communicate relevant information" (p. 108). As noted by the author, one of the key features that set indicators apart from other signs or forms of representation is the fact that indicators are relevant for decision-making (that is, judged relevant by decision-makers or the public). A similar argument is made by Lehtonen (2017) in his review of available definitions – the author notes that "[*T*]here appears to be broad consensus on at least one aspect of indicators: they are ultimately designed to be used" (p. 164)[55]

Building on Gallopin's (1996) definition, Bartl, Papilloud and Terracher-Lipinski (2019) elaborate on the key features typically associated with indicators as a policy tool, bringing forward the comprehensive character of the concept. The authors argue that indicators necessarily (*a*) represent a form of quantification (that is, they result from the process of assigning a numeric representation to the units of analysis); (*b*) entail a reduction in complexity (since any process of operationalization entails a reduction in the diversity of meanings ascribed to a given notion, and as indicators typically undergo a process of transformation during the production process); and (*c*) render certain phenomena visible

---

[55] Other scholars emphasize that what separates indicators from statistics is the fact that indicators are anchored in a conceptual framework, while statistics equate with raw data (see Gudmunsson, 2003). This criterium, however, has been characterized as problematic by Bartl, Papilloud and Terracher-Lipinski (2019), since no measurement exercise can be understood as a straightforward representation of reality, and any data-gathering effort rests upon a series of normative or analytical considerations of the nature of the researched phenomena, even if they remain implicit (as in the case of some statistic constructs).

(as they typically capture certain features that are not directly observable or which would otherwise go unnoticed). Remarkably, the authors make the case for relaxing the assumption that indicators are necessarily *relevant*. While relevance is one of the defining features typically emphasized by well-established definitions of indicators (including that of Gallopin, 1996), Bartl, Papilloud and Terracher-Lipinski (2019) observe that the relevance of a given statistic may be difficult to estimate – not least because its significance or visibility varies across social groups, and because its ultimate influence is highly contested.

The notion of an indicator thus encompasses a wide variety of simplified representations of the social reality. Different types of indicators can be organized along a theoretical-complexity continuum. For instance, Merry (2016) identifies three main types of indicators – borrowing (and expanding) Power's (1999) distinction between first- and second-order measurements. *Counts* constitute an initial modality of indicators oriented towards capturing the number or amount of a given unit or category of interest – with survey and census data featuring among the most prominent examples. *Ratio indicators*, in turn, compare two counts and thus entail additional interpretative work – as they require determination of not only the boundaries of the phenomena of interest but also the baseline of the comparison. Finally, *composite indicators* combine a variety of data sources into a single score and add an additional layer of complexity, for they require identification and weighting of those dimensions of a phenomenon judged to be relevant. Other categorizations classify indicators according to their potential for instrumental use or a policy function. Gudmunsson (2003) distinguishes between *descriptive indicators* oriented towards providing basic data, and *performance indicators* comparing a descriptive variable to a standard, target value or benchmark, and typically used for monitoring purposes.

Despite the diversity of numerical representations evoked by the notion of indicators, scholarship examining quantification as a tool of transnational governance appears to focus on a particular subset of indicators – specifically, those presented in a comparative format and/or designed for evaluation or auditing purposes (Bhuta et al., 2018). In some cases, this focus is implicit rather than explicit, and it stems from use of the indicator concept in a relatively narrow sense. This is, for instance, the case with Davis, Kingsbury and Merry's (2012) oft-cited definition:

> An indicator is a named collection of rank-ordered data that purports to represent the past or projected performance of different units. The data are generated through a process that simplifies raw data about a complex social phenomenon. The data, in this simplified and processed form, are capable of *being used to compare* particular units of analysis (such as countries, institutions, or corporations), synchronically or over time, and to *evaluate their performance* by reference to one or more standards. (p. 6; emphasis mine).

The centrality of the comparative and evaluative dimension is highlighted by Davis, Kingsbury and Merry's discussion of the defining features of the indicator concept (2012). These include (*a*) the significance of the indicator name – in that it plays a key

role in bringing certain phenomena into existence, that is, professing its factuality; (*b*) an ordinal infrastructure – that is, a hierarchical quality enabling comparisons over time or between different units; (*c*) an element of simplification or reductionism that translates complex phenomena into stylized, simple representations; and (*d*) the potential of indicators as tools for evaluation – in that they establish standards, and, either implicitly or explicitly, they establish and signal appropriate as well as pathological conducts.

Other well-known theorists focus explicitly on particular varieties of indicators that also emphasize the comparative dimension. Kelley and Simmons's work (2014), for instance, concentrates on *global performance indicators*, defined as "public, comparative and cross-national indicators that governmental, intergovernmental and/or private actors use regularly to attract attention to the relative performance of countries in a given policy area" (p. 4). The authors exclude explicitly those reporting systems that do not rate, rank or categorize countries. Similarly, Broome and Quirk (2015b) turn to the notion of *global benchmarking*, conceptualized as "an umbrella term for a wide range of comparative evaluation techniques that systematically assess the performance of actors, populations, or institutions" (p. 815). Similarly, Cooley's (2015) work zooms in on grading practices aimed at evaluating the performance of states – focusing on ratings (assigning values on the basis of pre-established performance standards) and rankings (oriented towards ordering a set of units and thus inherently relational and comparative).

Overall, and as the above suggests, literature on global indicators tends to privilege comparative and evaluative formats. This has important implications – particularly in that such emphasis has resulted in a certain neglect of other quantification practices of considerable significance for the study of global governance, including development of the international statistical system, the emergence of statistics on development, and UN statistical activity in particular (for an exception, see Ward, 2004). It is thus important to bear in mind that the theoretical propositions and analytical instruments outlined below should not and cannot be mechanically applied to any variety of global indicators, but are specific to comparative measures oriented towards assessing states' performance.

### 3.2.4. The impact of global indicators

Paralleling the rapidly increasing availability, use and visibility of global (comparative) indicators, a wave of recent scholarship has started to examine the appeal and effects of quantification endeavors. The theoretical apparatus developed by Merry remains one of the most influential contributions to this literature. In the seminal paper *Measuring the World: Indicators, Human Rights, and Global Governance* (2011), the author thus identified two analytically separate (albeit empirically interdependent) effects brought about by the expansion of indicators, namely knowledge effects and governance effects. This section builds on this distinction in order to give an overview of recent contributions to literature on the implications of global indicators for the practice of global governance.

<u>Knowledge effects</u>

Merry observes that indicators constitute, in the first place, a *technology of knowledge* in that they enhance the salience of certain issues, have the potential to redefine the meaning or contours of the concept they intend to measure, and can even bring into existence new objects – ultimately altering our perception of reality. The author notes that indicators shape actors' cognition and condition how problems and solutions are conceptualized, thus consolidating and naturalizing theories of social change that frequently remain implicit. The influence exerted by global indicators owes much to the appearance of objectivity and truth that they convey – and to the fact that they transform political processes into a series of (seemingly) technical issues (Merry, 2011, 2016; see also Davis, Kingsbury; and Merry, 2015).

Such observations resonate with a nominalist or constructivist approach to measurement in the sense advanced by Desrosières (2001). In the essay *How real are statistics? Four possible attitudes*, this perspective is defined in opposition to realist approaches to statistical labor. Realist perspectives[56] assume that measurement efforts simply capture social reality – and that measured objects have an existence independent of the act of measuring. A constructivist understanding rejects such propositions and is characterized by an explicit reflection on the constructed and negotiated character of statistical variables, and on the role of convention in coding and measurement processes. A constructivist (or nominalist) perspective contends thus that "[S]tatistical work not only reflects reality but, in a certain sense, establishes it by providing the players with a language to put reality on stage and act upon it" (Desrosières, 2001, p. 352). Indeed, more recently, Desrosières (2015) has explicitly cautioned against "careless" use of the measurement concept in that it tends to obscure the role of convention in quantification processes, as well as the creative dimension of such activity. Quantification would thus be made of two analytically separable moments – a moment of convention (fixing or revisiting meanings, negotiating categories and creating equivalences etc.) and a moment

---

[56] Desrosières (2001) identifies three different forms of realism – each one relying on different epistemological premises and characterized by a specific language and style of reasoning. *Metrological realism* takes inspiration from the theories of measurement advanced by natural sciences, thus assuming that "[T]he object to be measured is just as real as a physical object" (2001, p. 341). Metrological realism is thus centrally concerned with the question of reliability and emphasizes bias and measurement error as the main obstacles to "real measures" (Desrosières, 2001; 2015). *Accounting realism* or *pragmatic realism* is inspired by business accounting practices and derives certainty from the consistency and reproducibility of evaluation and measurement practices. Since business accounting requires estimation of the monetary value of a series of items that are not directly observable (e.g., expected gains and the value of doubtful loans etc.), the realness of these objects results from the trust inspired by numbers – in turn, a product of the application of consistent accounting procedures. Finally, the *proof-in-use approach* captures the reality judgments of those data users who are not engaged in data production and who use statistics for argumentative purposes. Such users are primarily concerned with the plausibility and consistency of the results derived from a given dataset, whereas they have little interest in the data-production process (Desrosières, 2001; see also the discussion in Espeland and Stevens, 2008).

of measurement (that is, the act of counting, registering and assigning values on the basis of such conventions).

When it comes to global indicators, and as noted by Merry (2011), knowledge effects are particularly obvious in the construction of *universal* categories which, in order to render the world intelligible, necessarily erase contextual particularities – thus creating a new form of knowledge and theoretical constructs amenable to cross-country comparisons. A similar argument is made by Broome and Quirk (2015a), who observe that the processes of numerical translation behind global benchmarking and the construction of universal constructs typically reduce complexity and contextual detail, engaging in processes of simplification and reification, thus producing a sense of homogeneity among otherwise heterogeneous objects.

Knowledge effects are also especially visible in those cases where the production of new indicators displaces other metrics – as such processes typically lead to a redefinition of the phenomenon. The Human Development Index (HDI) offers an illustrative example of these dynamics. As documented by Stanton (2007), the HDI was developed in order to counter the prominent role of GDP as the primary measure of development and aggregate social welfare. First published in 1990, the HDI indeed succeeded in advancing and popularizing a new conceptualization of well-being – one that drew heavily on the capabilities approach developed by Sen (1985) and which emphasized ends (effective access to resources and a decent standard of living) over means (income per capita). Such a shift had important implications – particularly as the HDI challenged the well-established assumption that the ultimate goal of development was to increase national income, and advanced a number of new concerns within the development realm.

Governance effects

A second dimension of impact identified by Merry (2011; 2016) is that of *governance effects*, sometimes referred to as *power effects* (Davis et al., 2015). Such notions bring forward the potential of indicators to transform the power dynamics behind decision-making procedures, as well as their regulatory effects on the behavior of a variety of actors. The use of indicators as a technology of governance is sometimes relatively obvious – for instance, when they are used to inform decision-making regarding the allocation of aid, foreign investment or other resources. Other governance effects, however, are less evident – including consolidation of power in the hands of those exhibiting expert authority, and promotion of self-governance among monitored subjects (Merry, 2011).

These questions have been explored in a more systematic way by Davis, Kingsbury and Merry (2012), who identify four possible effects of indicators on global governance. The authors observe that, in the first place, indicators "can be expected to affect where, by whom, and in relation to whom governance takes place" (p. 12), thus having direct

74

implications for what the authors term the *topology of global governance*. More specifically, the authors argue that indicators are likely to confer some degree of authority to actors that would otherwise wield little influence (including the promulgators but also the producers of indicators), and also shape and constitute new identities, bringing together previously disconnected organizations, and structuring highly decentralized or informal governance structures.

Secondly, indicators transform the processes through which standards and norms are set. Insofar as indicators are used with evaluative purposes, they work as placeholders, carriers or markers for theories and policy ideas – even if these are not necessarily made explicit. The significance of indicators as *vehicles for standard setting* lies in the fact that, unlike other politically explicit norm-building exercises, the production of indicators rests typically upon scientific debate and shifts the balance of power towards scientific expertise. The specificity of indicators as standard-setting instruments is synthesized by the authors' observation of the fact that:

> Whereas political efforts to formulate norms and standards—for example, in multilateral intergovernmental negotiations conducted by diplomats—tend to involve processes such as voting, interest-group bargaining, or the exercise of material power, the processes in specialist agencies and expert meetings at which the standards embedded in indicators are produced, accepted, and supported tend to involve derivation of power from scientific knowledge. (Davis et al., 2012, p. 83)

Thirdly, indicators affect the *practice of decision-making* among a wide range of actors. Importantly, the regulatory effect of indicators does not result exclusively from their use among powerful organizations who use them to "nudge" states in the right direction or to inform resource allocation. Davis, Kingsbury and Merry (2012) note that indicators can transform the behavior of targeted states as they are likely to eventually become internalized as "guides to appropriate conduct" among domestic constituencies and decision-makers, but also because indicators are frequently mobilized to justify or legitimize local decision-making processes, in virtue of their aura of objectivity (i.e., efficiency, scientific soundness, transparency and impartiality etc.).

It should be noted that questions relative to the regulatory impact of indicators are frequently discussed with particular reference to ranking practices. Thus, a number of authors have examined the effects of quantification on decision-making processes by focusing on the *comparative* dimension of indicators. This is the case with Broome and Quirk's (2015a) discussion of the political effects of global benchmarking. The authors note that such effects "tend to be cumulative and subtle, rather than overt and immediate" (p. 826) and that the influence of such comparative metrics rests upon their capacity to affect how actors conceptualize their opportunities, obligations and policy options, and how they justify their relative performance. Borrowing from Barnett and Duvall's (2005)

notion of productive power,[57] Broome and Quirk (2015a) emphasize the *indirect* nature of the power exerted by benchmarking practices. According to the authors, the potential of such instruments lies primarily in their capacity to stimulate and shape a series of political conversations concerning the salience and priority of certain themes; the performance of a given actor vis-à-vis its peers; and the evolution of such performance over time. Hence, rankings tend to have a validating effect among those scoring highly – for whom they serve as a means to preserve and justify the status quo. Conversely, a bad performance is likely to be instrumentalized among those seeking a policy change, and to stimulate a desire for reform among decision-makers (frequently motivated by fear of material but also social sanctions, including name-and-shame dynamics and peer pressure). Through these dynamics, global indicators contribute to the reification of normative standards, as well as to consolidation of a certain normative consensus. In a more recent article, Broome, Homolar and Kranke (2018) go on to argue that benchmarks constitute a key source of power for IOs in that they enable them to identify and disseminate ideal and pathological models of state behavior. Benchmarks allow IOs to act not only as teachers and judges of norm compliance but also as evaluators of countries' performance, thus contributing to reinforcing their normative power. As noted by the authors:

> By holding up some states as role models to emulate while framing others as underperformers who have to change, IO benchmarks promote images of the world as divided into cases of "success" and "failure". By linking the attribution of praise and blame to knowledge-based practices, IO benchmarking specifies "what is normal and desirable" and, by implication, "what is abnormal and undesirable behavior" (p. 531)

A similar argument is raised by Kelley and Simmons's (2019) work on global performance indicators. The authors argue that the power of such indicators stems from their capacity to affect reputations, noting that "What makes this form of information deployment potentially powerful is its ability to broadly impact social beliefs about successful states and appropriate policies. It is deployed precisely as a form of social pressure on targets to conform – or suffer the reputational consequences, real or imagined" (p. 498). A key contribution by these two authors is their systematization of the mechanisms through which global performance indicators influence policy choices (Kelley and Simmons, 2014). These include the *transformation of domestic politics* (frequently mediated by the mobilization of domestic coalitions that invoke rankings in order to press for change); *direct elite responses* (if indicators stimulate a transformation of organizational routines within government or administrative institutions, transforming the priorities of government officials and top bureaucrats[58]); and *transnational pressures*

---

[57] Defined as the "socially diffuse production of subjectivity in systems of meaning and signification" (Barnett and Duvall, 2005, p. 43).

[58] This might be the result of genuine internalization of norms embedded in the indicator (frequently the product of intensive consultation processes and long periods of dialogue taking place before or after release of the ranking), but it might also be the result of a more strategic

(when the publication of market expectation indicators for a given country inspire and condition the decisions of private economic agents, or inspire third parties to exert some form of pressure over the rated countries).

Returning to Davis, Kingsbury and Merry's (2012) categorization, in the fourth and last place, indicators are found to affect *contestation dynamics*. Even if indicators may complicate contestation (in that they tend to obscure the political theories that inspire them), uncritical acceptance is uncommon or short-lived at best. Thus, the more significance a given indicator acquires, the greater the chances it will become the subject of contestation and debate. While disputes over indicators can take conventional forms (including lobbying and litigation), they can also give rise to new strategies, ranging from opposition to data-collection efforts, to the creation of alternative indicators or engagement in scientific disputes. In connection with such dynamics, the authors observe that the use of indicators as a technology of governance is likely to trigger efforts to regulate the production and use of indicators. This could be the case with those initiatives aiming to subject the producers of indicators to strict procedural standards, or to educate the public on both the potential and risks of quantitative measures (Davis et al., 2012). Thus, while production of indicators relies on scientific debate and a shift in the balance of power towards scientific expertise, this is likely to change as the normative effects of indicators are rendered more evident. As noted by the authors:

> As the awareness or the significance of indicators as standards rises, indicator design and production are likely to become increasingly subject to demands made of other standard-setting processes, including demands for transparency, participation, explanation, justification, and review. (2012, p. 83)

Such observations are particularly relevant since they sound a note of caution against a totalizing understanding of the governance effects of indicators – one oriented by what has been termed "neo-Foucauldian tendencies" (cf. Bartl et al., 2019), that is, by an excessive emphasis on the power of indicators, which tends to neglect the agency exerted by monitored actors. Such approaches fail to account for episodes of resistance and, more generally, fall short in terms of explaining the recontextualization and translation of indicators at the local level (Bhuta et al., 2018; Erkkilä and Piironen, 2018). The limited attention given to such vernacularization processes leads, in turn, to an inadequate understanding of the ways in which local actors negotiate the meaning of indicators and engage in their reinterpretation, or put them at the service of domestic and pre-established agendas.

An associated problem with these Foucauldian perspectives is the question of the *intentionality* behind production of indicators. Erkkilä and Piironen (2018) have noted that the rationales guiding production of indicators are more complex than most accounts suggest – noting critically that the production of indicators is unlikely to be driven by a

---

response in which decision-makers review policy advice produced by the monitoring agency, or even seek advice from the producer of the indicator (Kelley and Simmons, 2014).

single, rational and well-defined rationale or motivation. The question has been examined more extensively by Dahler-Larsen (2013) in a discussion of the constitutive effects of performance indicators. According to the author, much of the literature on so-called "unintended effects" of numbers is guided by problematic assumptions regarding the intentions of those governing production of numbers. Thus, the author notes:

> More often than not, it is just assumed that there was a point in time where an architect behind a given stable and bounded indicator system had a consistent and coherent vision that defined one agreed primary objective for the indicator system against which all of its consequences can be judged. (Dahler-Larsen, 2013, p. 975)

A final clarification should be made in relation to the governance effects of quantification practices. While Davis, Kingsbury and Merry's (2012) categorization (as well as much of the literature discussed above) focuses on empirically observable effects of indicators, other scholars have taken a more theoretically oriented approach – examining how these governing effects are possible in the first place, that is, determining the properties of indicators that make such governance effects possible. A particularly comprehensive example of such a line of inquiry is Erkkilä and Piironen's (2018) categorization of these enabling mechanisms. Weaving together the theoretical insights advanced by early theorists of quantification, the authors argue that the governing effect exhibited by numbers results from four different mechanisms, namely: *objectification* (the process through which a given notion becomes a collectively shared notion); *(de)politicization* (i.e., the potential of numbers to broaden and narrow the scope for debate, and to selectively articulate some issues as policy problems); *subjectification* (the process through which certain agents acquire an identity that shapes their behavior as well as the expectations placed upon them, and through which evaluated units come to be understood as autonomous and therefore responsible); and *legitimation* (i.e., the possibility to confer an aura of objectivity and credibility to certain agents, generally associated with expert authority). While such questions are of limited interest for the purposes of this dissertation and are not pursued further, it is necessary to clarify that they represent an alternative approach to the perspective privileged by much of the foregoing literature.


*3.2.5. The case for organizational effects*

Much of the literature on the implications of global indicators has tended to focus on their influence over monitored populations, the users of indicators or a loosely defined (global) public. Conversely, and as noted by Freistein (2016), the effects of indicators on the organizations producing them have been less systematically examined. The author brings to the fore what she terms indicators' *organizational effects* – that is, their impact on the operational logic of IOs. The author thus notes that creation of indicators can be understood as a ritual motivated by concerns other than efficiency or problem adequacy – and in particular, by the need to signal and ascertain authority over certain issues. Thus, the creation of indicators is oriented towards affecting the organization's environment but also its own identity. Freistein (2016) therefore notes that indicators can be used to

communicate *within organizations* (encouraging the development of a shared identity but also enabling coordination among different units operating within an IO) or *across organizations* (fostering relationships of competition or cooperation). She also notes that indicators have a social function as markers of competency and authority, and tend to establish path-dependent or self-generative logics – so that, at some point, organizations cannot afford to not engage in quantification efforts.

While Freistein's (2016) observations are among the few that bring organizational effects to the center of the analysis, it should be noted that much of the literature addressing the supply side of the indicator market tacitly assumes that indicators have the potential to transform the status, identity or visibility of the responsible organizations. In their discussion of indicators' effect on the topology of governance (see the section *Governance effects* above), Davis, Kingsbury and Merry (2012) noted that one of the key effects of indicator construction is the possibility of enhanced organizational status or visibility and also the transformation of inter-organizational relationships – as the production of an indicator brings together and brings into contact a range of disparate actors who partake in its construction or are concerned with the phenomena it intends to capture.

There is, in fact, growing recognition that organizations engage in the production of indicators as a strategy to position and (re)brand themselves vis-à-vis other institutions. Cooley (2015), for instance, argues that rankings and rating are used by both governmental and non-governmental IOs as a means to assert their authority over specific issues, and to posit themselves as central organizations in this area. Ranking and rating are thus conceptualized as organizational branding (or "flag-planting") exercises, as the production of such devices "can be an invaluable tool in staking the organization's claim to govern that issue and advance the solutions to the problem in question" (p. 21). According to Cooley (2015), the use of indicators as branding devices is especially relevant for think tanks and policy institutes, as well as for NGOs or other advocacy organizations vying and competing for visibility, credibility or resources in an increasingly saturated market. An example of such dynamics would be the case of the Corruption Perception Index, which would have bolstered the public profile and reputation of Transparency International (Cooley, 2015). Broome and Quirk (2015a) introduce a nuance to this argument, noting that the potential of benchmarking practices as sources of organizational prestige is mediated by their capacity to project an image of technical expertise – understood as an increasingly valued source of authority in transnational governance dynamics.

Whereas these accounts tend to focus on inter-organizational competition, Arndt (2008) draws attention to the impact of quantification labor on intra-organizational dynamics. In her discussion of the World Bank governance indicators, the author notes that institutional reasons are one of the key drivers behind the rise of governance indicators, and observes that production of indicators constitutes a means to raise the visibility of a given IO vis-à-vis its competitors, but also as a way to enhance the status of particular units or divisions

*within* a given organization. This would, for instance, be the case of the International Finance Corporation or the World Bank Institute, whose status within the World Bank was greatly enhanced partially as a result of the traction gained by their flagship indicators – the Worldwide Governance Indicators and the Doing Business indicators, respectively (Arndt, 2008).

*3.2.6. The making of an indicator: theoretical considerations*

In their comprehensive overview of recent trends within the so-called sociology of quantification, Diaz-Bone and Didier (2016) note that it is necessary to go beyond analysis of the uses and impacts of quantification,[59] and pay greater attention to the processes of quantification itself. A similar argument is made by Davis, Kingsbury and Merry (2012) in their seminal work on global metrics. The authors argue that a proper analysis of the effects of indicators requires an understanding of the process of "development and crystallization of indicators over time" (p. 10). To this end, and borrowing from Halliday and Shaffer's (2015) notion of normative settling, the authors have, more recently, proposed a distinction between four different phases into which the trajectory of an indicator can be operationalized[60] (Davis et al., 2015). These include the conceptualization of indicators, their production, their use and their influence. This section describes the most distinctive properties of these different phases and examines them in relation to theoretical developments specific to each stage.

Conceptualization

The *conceptualization* of indicators captures the process through which a given indicator is named and formulated – including the definition of categories but also the criteria and analytical procedures for measurement. As noted by the authors, this process typically builds on a theory of sorts, although this is likely to remain implicit. This is particularly the case for those indicators aimed at capturing complex phenomena (and thus relying on multiple measures), for the selection of these relevant subdimensions entails a theory of causation, establishing the determinants of a given phenomenon. One of the potential by-products (or unintended consequences) of this first phase is the so-called "indicator creep", that is, "the gradual expansion of obligations as new measurements are created to

---

[59] The term is used by the authors to capture those analyses of quantification concerned with the social and political dimension of these phenomena. The authors make explicit that, despite its labeling, this emerging field is not restricted to sociological approaches but constitutes an essentially transdisciplinary space (Diaz-Bone and Didier, 2016).

[60] The authors note that although the four stages are presented as discrete categories, this is an essentially theoretical separation; in practice, the different phases are likely to mutually affect one another, particularly if the indicator is revised or modified.

operationalize standards that extend beyond the initial standards" (Davis et al., 2015, p. 11).

Another key point raised by Davis, Kingsbury and Merry (2015) in relation to this first stage is the fact that there is no mechanical connection between the theory informing a given indicator and the ideological orientation of the agency promulgating or producing it. Coherence depends ultimately on a variety of factors, including the degree of normative or ideational convergence exhibited by the institution and the relative significance of the indicator for the organization.


Production

The second stage in an indicator trajectory is that of *production* – that is, the phase in which the conceptualization of the indicator is coupled to an already existing or newly created dataset. One of the key components of the production stage is the *promulgation* of the indicator, that is, its packaging and dissemination. However, the authors caution against placing excessive emphasis on the promulgators of a given indicator – especially if it comes at the expense of other relevant agents involved in its production. In fact, one of the most relevant points raised by the authors in relation to this stage is the fact that the production of global indicators typically relies on the data-collection efforts of a wide (and frequently disperse) range of actors. This is likely to translate into a number of challenges associated with the difficulty of securing the cooperation and diligence of data suppliers, which might under-report, over-report or reinterprete categories; as well as with contestation dynamics between different data sources (most notably, between national and international suppliers).

This collaborative nature of the indicator-production process is of particular relevance. As noted by the authors in previous works, one of the factors that complicates proper understanding of the effects of indicators on their own creator is the fact that the production of a global indicator is a *collective* endeavor that goes beyond its main promulgator (Davis et al., 2012). Rather, this process requires the data-collection efforts of a range of actors (including IOs, national statistical offices and NGOs etc.), and needs to incorporate the methodological insights advanced by scientific or scholarly communities. As noted by the authors:

> [*T*]he promulgator of an indicator may or may not be the actor most involved in determining its content. Instead, the promulgator is often more like a consumer-product manufacturer, whose main contribution is to lend its brand name, and perhaps its design and marketing expertise and quality-control power, to the collective product of a global supply chain. (2012, p. 82)

However, limited attention has been paid to the social labor behind this transformation of raw data into authoritative representations. Key contributions to these questions have been advanced by Espeland and Stevens (2008). The authors draw attention to the fact

that the authority or certainty of a given indicator echoes the status of its handlers or users – or, to put it differently, indicators become reliable as they "move upwards":

> "Raw" information typically is collected and compiled by workers near the bottom of organizational hierarchies; but as it is manipulated, parsed, and moved upward, it is transformed so as to make it accessible and amenable for those near the top, who make the big decision. This "editing" removes assumptions, discretion and ambiguity, a process that results in "uncertainty absorption": information appears more robust than it actually is [...] An often unintended effect of this phenomenon is numbers that appear more authoritative as they move up a chain of command. The authority of the information parallels the authority of its handlers in the hierarchy. (Espeland and Sanders, 2008, p. 421–22)

There is, in fact, evidence that the producers of indicators are aware of the imperfection of their own datasets but do not see this approximate character as inherently problematic. Rocha de Siqueira (2017) has put forward the notion of "good enough data" in order to shed light on these dynamics. The concept captures the specificity of data recognized by both producers and consumers as imperfect (or even prone to error) but accepted as an authoritative source in virtue of their practicality or convenience. Overall, the crucial transformation of "raw" data, and particularly their conversion into authoritative numbers despite the limitations, calls for greater attention to the coordination efforts and social and intellectual labor entailed by the production process.

Use of indicators

The production of an indicator is followed by a third stage corresponding to *the use of indicators* (Davis et al., 2012). This phase captures the different activities and procedures through which a given indicator is mobilized as a source of knowledge – to form or support beliefs, guide and inform research, and orient decision-making etc. Importantly, there is no direct correspondence between the intentions and expectations of those in charge of the conceptualization and production of indicators, and the final uses eventually given to them. This is the case as the use of given statistics results from the activity of a diversity of stakeholders over which the producers of indicators have only limited and indirect influence – including national or international bureaucrats and officials, NGOs, civil society, the research community and the general public.

It is, however, important to note that the ultimate use given to an indicator is at least partially mediated by its *utilization framework* – that is, the nature of the accountability mechanisms associated with the metric. Gudmunsson (2003) distinguishes between three main varieties of utilization or accountability frameworks, namely: (*a*) information frameworks, in which indicators are simply communicated to the general public with the expectation that the information will be "taken into consideration" – in an undefined way; (*b*) monitoring frameworks, which include regular reporting provisions and feedback mechanisms, in the expectation that (policy) action will be taken if the indicators suggest policy failure; and (c) control frameworks explicitly oriented towards regulating policy-

making by way of comparing the results against pre-established standards or benchmarks and assigning responsibilities to a specific set of agents.

Thus, utilization frameworks associated with indicators frequently entail the identification and distribution of responsibilities. When it comes to global benchmarking exercises, this attribution of responsibilities presents some particularities, as advanced by Clegg (2015). The author contends that the relationship between benchmarkers and benchmarkees is a complex one in that it is not as hierarchical, fixed or unidirectional as these labels might suggest. Clegg (2015) argues that the lines of responsibility associated with global benchmarking are the result of two separate moments – the design phase and the implementation phase. In each of these two moments, responsibilities can be assigned in either direct or diffuse terms. This leads to a fourfold typology of benchmarking modes which capture both the original delineation of responsibility lines and their evolution over time. One possibility is *direct responsibility*, in which the architects of the benchmark clearly identify a set of agents as responsible for the outcome – and these are effectively subject to material and/or reputational sanctions for underperformance. The second possibility, *diffuse responsibility*, captures the opposite situation – in which the *locus* of responsibility is not identified during the benchmarking design phase and remains unspecified over the lifetime of the benchmark (as a consequence of a lack of consensus or given the limited authority of the benchmark's architects). A third possibility is that of *strategic blurring* – when those originally identified as responsible succeed in minimizing both reputational or material sanctions by adducing mitigating factors or the responsibility of third actors – thus engaging in a "blame game". The fourth possibility, labeled *strategic clarification*, captures the opposite sequence. It includes those instances in which the lines of responsibility emerge only with the passage of time. This might be the result of either push-related dynamics (when third-party actors succeed in clarifying the lines of responsibility, frequently in a context of goal failure) or pull-related dynamics (when a benchmarked agent succeeds in positing itself as responsible for its position in the ranking, frequently in the context of goal success).

Impact and effects

The fourth and last stage in the trajectory of an indicator corresponds to its *impact and effects*, that is, its ultimate influence – which does not necessarily correspond to its use(s) (Davis et al., 2015). The need to distinguish between the use of an indicator and its effects has been highlighted by Lehtonen, Sébastien and Bauler (2016), who note that while the notion of use conveys an element of intentionality and awareness, an indicator can influence a given policy area in more indirect ways – for instance, by changing the terms of the debate or tying together previously unconnected actors. These authors thus note that there is no direct correspondence between use and influence, given that "[*F*]requent use of indicators does not guarantee that they would be highly influential, whereas even indicators that are not actively used may nevertheless generate significant impacts through indirect pathways" (p. 4).

The impact of indicators is the object of an expanding corpus of scholarship – including literature on so-called knowledge and governance effects, discussed in section 3.2.4. However, Davis, Kingsbury and Merry (2015) warn of the difficulties in assessing the magnitude of such impact, particularly given the difficulties in isolating the effect of indicators from a range of confounding factors. In addition, further research is needed to understand the determinants of this influence – that is, the factors explaining why some indicators gain traction while others do not, or possible reasons behind the decline of certain metrics. Some possibilities advanced by the authors include the prestige of the promulgator, the indicator quality, the indicator ecology, and supply-and-demand dynamics (Davis et al., 2015).

This last question has been examined in great depth by Büthe's (2012) political-economic model of indicators, which distinguishes four distinct sets of stakeholders relevant to understanding production of indicators, going beyond simplistic supply-and-demand models. Büthe firstly identifies the *rule-demanders* group, which actively calls for measurement efforts, sometimes to the point of being willing to provide financial support. This group plays a crucial role in nurturing and encouraging the demand for indicators, even if the group is not necessarily in charge of production. A second group is that of *rule-makers*, who "write, maintain and disseminate such measurement standards" (2012, p. 32). The third group is constituted by indicator users, that is, those relying on a given metric in their own decision-making procedures. Finally, the fourth group corresponds to a subset of users – the targets of indicators, that is, those directly measured by quantification efforts. Büthe (2012) notes that while, on occasion, some of these groups can overlap, this is not necessarily the case. The author argues that the examination of possible divergences between these groups, as well as the interrogation of the power dynamics that govern their relationships, is a necessary step in order to adequately understand the role of indicators in global governance. For the purposes of this dissertation, Büthe's (2012) distinction is of particular relevance in that it complements and expands on Davis, Kingsbury and Merry's (2012) considerations of the collective nature of quantification labor – and the need to go beyond the most visible or high-profile agents – namely the promulgators of indicators.


## 3.3. Making sense of organizational independence

The third and last theme addressed in this chapter is the question of inter-organizational interplay. This theoretical problem is of particular relevance given the dissertation's interest in interorganizational negotiations during the run-up to adoption of SDG4, as well as different forms of organizational interdependence brought about by the new monitoring needs. Because SDG4 was born out of multiorganizational dynamics, and its monitoring and implementation weaves together different organizations, it is necessary to identify those theoretical constructs better suited to capturing inter-organizational interaction – but also to make sense of the multilevel character of such processes.

To this end, the dissertation turns to *field theory*. This section discusses the theoretical concerns that inform emergence and consolidation of this approach within IR literatures, as well as the assumptions that inform it. It elaborates on the properties of *transnational fields of governance* as a notion specifically developed for the study of soft regulatory processes, and it discusses those dimensions of analysis identified as being most relevant to charting and inquiring into the structure of transnational fields.

### 3.3.1. Towards a relational sensibility

An initial difficulty associated with the analysis of multi-actor and inter-organizational dynamics such as those shaping SDG4 stems from the fact that the question of inter-organizational relations (IOR) is the object of a rather reduced and fragmentary body of IR scholarship. As Biermann and Koops (2016) observed in their overview of the so-called inter-*organizational* turn within IR, it was not until the early 2000s that theoretically driven scholarship started to grapple with the interaction between IOs. This wave of scholarship emerged in the light of proliferation of both international governmental and non-governmental organizations, as well as their increasing interaction – in turn, connected to the aggrandizement of their own mandates and missions, and the emergence of global agendas tying together previously disjointed fields. However, and despite growing empirical interest in the interaction between IOs, IR approaches have tended to lag behind the theoretical advances made by other disciplines, including sociology and public administration sciences (Biermann and Koops, 2016). Constructivist perspectives in particular have tended to rely on *implicit* theories of IOR (Lipson, 2016).[61] More specifically, research on epistemic communities and the bureaucratic and professional cultures of IOs has tacitly put forward certain theses on the potential drivers, and enabling and constraining factors behind IOR – frequently highlighting the role of shared ideas and beliefs. More generally, IR has typically borrowed from organizational theories in order to make sense of the interaction among IOs. A repertoire of theories advanced by organizational studies has thus been imported into the IR field in order to compensate for such a void – including principal-agent, resource-dependence and network theory. However, and as noted by Lipson (2016), the theoretical potential of such perspectives for the study of IOR remains largely untapped.

---

[61] To be sure, the most theoretically sophisticated accounts of IOR (grounded in a distinctly IR perspective) have adopted a rationalist approach, frequently turning to *regime theory*. This has particularly been the case of the study of environmental cooperation, which would have taken the lead in analysis of the interplay among international organizations (see section 2 in Lavenex, 2014, for a brief overview). International regimes are conventionally defined as "sets of implicit or explicit principles, norms, rules, and decision-making procedures around which actors' expectations converge in a given area of international relations" (Krasner, 1982, p. 186). Prominent regime-theory scholars have addressed the question of organizational interplay, conceptualizing different forms of cooperation (including interorganizational cooperation) as instrumentally oriented decisions judged appropriate for the pursuit of agents' interests (Lipson, 2016).

An associated difficulty faced by IR scholarship on IO interdependence results from the fact that, as observed by Seabrooke and Sending (2015), both constructivist and rationalist approaches have tended to conceptualize IOs as actors with a set of predefined and fixed attributes – thus neglecting the constitutive effect of interaction. On different occasions, Sending and Seabrooke (along with their collaborators) have voiced concern over a conceptualization of such entities as self-contained and sharply bounded spheres of action with a fixed set of attributes, and have advocated for a more complex understanding of these actors, driven by the assumption that their defining features are permanently negotiated through their interaction with other entities. They also express concern over IR's tendency to "black box" IOs and also argue that IOs' boundaries are more porous or permeable than is frequently assumed (Seabrooke and Sending, 2015; Sending and Neumann, 2011).

Such observations need, in turn, to be understood in relation to broader concerns with IR's tendency towards theoretical reification (Seabrooke and Sending, 2015). In recent years, different scholars have taken issue with the reified approach to both state and non-state actors that has long characterized IR theory – that is, the reliance on an excessively abstract and fixed understanding of such entities (Adler-Nissen, 2013; McCourt, 2016). These authors have been particularly critical of actor-centric theorizations that tend to overlook the key role played by interaction in the very constitution of identities, interests or preferences. Borrowing from Emirbayer's (1997) distinction between substantialist and relational social theory, they have noted that most traditions within IR fall within the former in that they assume that the entities of interest (the state but also other actors) have an existence of their own, prior to their interaction with other units (Adler-Nissen, 2015; Jackson and Nexon, 1999). A relationalist approach contends instead that relations need to be seen "as constitutive and ontologically prior to these entities" (Sending, Pouliot and Neumann, 2015).

### 3.3.2. Bringing field theory into IR

In response to the theoretical problems raised by the relationalist propositions discussed above, a new line of inquiry has unfolded over the last decade, driven by the objective of advancing towards a less essentialist and more situated understanding of actors operating at a transnational level. Efforts to bring an explicitly relational sensibility to the study of IR and transnational governance, as well as to pay greater attention to the role of practical knowledge, have thus crystallized into a growing interest in Bourdieu-inspired *field theory.*[62] Field theory as imported into IR is thus explicitly aimed at countering and

---

[62] Although Bourdieu features as one of the most prominent or best-known proponents of field theory, it should be noted that the field notion (or metaphor) has been used in a different sense by other theorists – see, for instance, Fligstein and McAdam's (2012) *A Theory of Fields*, as well as DiMaggio and Powell (1983). However, in the interest of conceptual clarity, and as discussed in

compensating for the reified or excessively abstract accounts perceived to be characteristic of much IR scholarship (Adler-Nissen, 2013; Cohen, 2018). Through its emphasis on the explanatory power of actors' relative positions, field theory is expected to succeed in overcoming the reification and substantialist tendencies that would (allegedly) characterize a substantive part of IR scholarship. According to this author, field theory is perceived as a way out of the structuralist bias of certain constructivist accounts which emphasize collective identities and preferences, but come short of explaining how actors were socialized into these dispositions.

Thus, even if, as discussed below, fields operating at a global level are likely to exhibit moderate autonomy at best, the potential of field theory lies in its capacity to advance a relational approach to the study of transnational dynamics. This perspective brings to the fore relationships among actors as key shaping forces modeling these entities. This way, it avoids portraying IOs (and other actors) as if operating in a vacuum. It makes it explicit that their interests or behavioral proclivities do not result from qualities intrinsic to the organizations but from their relative position and the context in which they are embedded. The added value of this approach is summarized by Pouliot and Mérand (2013) in their overview of the uses of a Bourdieusian perspective within the field of IR:

> To think in terms of fields […] is to think in terms of relations. If we apply [*Bourdieu's*] work to IR, this relational approach allows us to recognize a level of analysis that is quite distinct from the discipline's dominant currents: it is not focused on substances, such as the state and state actors, or essentialized concepts such as politics or globalization, but instead on the "totality of relations" involving the positions that are uncovered, structured, and conceptualized in the field. (p. 32)

More generally, the rise of field theory within IR needs to be understood in connection with the recent interest on Bourdieu's theoretical and empirical toolbox, and in the application of Bourdieu's political sociology to transnational and international phenomena. This interest has been driven by the assumption that certain recurrent themes in IR theory could benefit from a Bourdieu-grounded analysis, but also by the expectation that Bourdieu's work may help to shed light on previously neglected or undertheorized aspects (Cohen, 2018). Bourdieu-inspired theory is thus portrayed as a means to bring power to the fore, drawing attention to the theoretical centrality of the struggles for power, prestige or authority – a theme sometimes overshadowed by the constructivist emphasis on intersubjectivity (Guzzini, 2013).

Additionally, Bourdieu's constructs (including field but also habitus or capital) are conducive to securing greater attention to the micro-sociological foundations of international life – by placing at the center of the analysis the constitutive impact of those routines and habits perceived as commonsensical (McCourt, 2016). In this sense, Bourdieu's work has been considered instrumental in advancing and consolidating the

---

3.3.3 (*Reconceptualizing the notion of field*), this dissertation focuses on the notion of field in the sense advanced by Bourdieu (1993).

so-called *practice turn* within IR (Adler-Nissen, 2013). The practice crystallized in IR theory during the late 1990s and 2000s, mirroring, in turn, a broader shift in social sciences championed by Shatzki, Knorr Cetina and von Savigny's *The Practice Turn in Contemporary Theory* (2001), and Reckwitz's *Towards a Theory of Social Practices* (2002). Despite its theoretical pluralism, practice theory can be characterized by its focus on tangible, observable instances of human activity (that is, practices) as key units of analysis. Another key contribution of this approach has been the problematization of the logic of consequences and the logic of appropriateness as key drivers of social action – privileged respectively by rationalist and constructivist accounts (Cornut, 2017). Practice theory instead draws attention to the role of practical sense and habit as key forces shaping human action – noting that "[*E*]ven while an agent is reflecting upon what action will yield the most benefits or correspond to her normative commitments, she is doing so against a taken-for-granted background of habit that has already constrained her imaginable outcomes" (Hopf, 2010, p. 547).

Specifically in the field of IR, the practice turn was introduced by the seminal works of Neumann (2002), and Adler and Pouliot (2011), who drew extensively on Bourdieu's work to substantiate their claims and articulate their theoretical apparatus. The practice turn as advanced by such works calls for greater attention to the role of practical knowledge and "irreflexive" (that is, pre-intentional) practices in the structuration of world politics and transnational phenomena (Pouliot and Mérand, 2013).[63] IR scholars situating their work within this perspective coincide in their emphasis on the more "mundane" unfolding of transnational social action, and particularly, on a certain skepticism regarding the explanatory power of highly abstract constructs. The practice turn in IR favors instead greater empirical attention to the practices, routines and worldviews exhibited by transnational actors – thus contributing to countering their reification.

### 3.3.3. Reconceptualizing the notion of field

Against this background, the concept of field has been gaining currency as a productive analytic device which can serve to promote the understanding of a number of internationalized subject areas. Even if empirical studies inspired by such an approach have tended to remain restricted to a few issues (including law, security, culture and European integration) (cf. Cohen, 2018 for an overview; see also Madsen, 2006; and Pouliot and Mérand, 2013), field theory is perceived as being particularly apposite to the study of transnational interactions and the governance of transnational spaces and, more generally, to the theorization of transnational spaces as social spaces (Christensen, 2017).

---

[63] To be sure, practice theory in IR constitutes a particularly fragmented space and accommodates a diversity of perspectives that are beyond the scope of this research. See Kustermans (2016) or Martin-Mazé (2017) for a comprehensive account of the different perspectives informing this approach.

At its simplest, the notion of social field as defined by Bourdieu refers to a "structured space of positions (or posts) whose properties depend on their position within these spaces and which can be analysed independently of the characteristics of their occupants (which are partly determined by them" (Bourdieu, 1993, p. 72). Defining features of social fields include (*a*) power relations – as fields are made of a number of unequal positions determined by uneven access to different forms of capital, which consequently give rise to a hierarchy of domination; (*b*) a series of stakes (or objects of struggle) specific to the field and over which different agents compete; and (*c*) a series of principles and rules observed and shared by all agents in the field (encompassing ideas, norms and knowledge) and a shared belief in the relevance of certain stakes (Leander, 2010; Swartz, 2013).

However, importing the notion of field at the global level is not without problems and requires a re-conceptualization effort. This is first and foremost a consequence of the fact that the notion of field remains a "fuzzy" and contested term, for which Bourdieu provided multiple (and sometimes contradictory) definitions (Leander, 2010; Martin, 2003). In addition, Bourdieu himself never examined IR categories or themes in much depth as his works focused almost invariably on the (French) domestic arena (Cohen, 2018; Jackson, 2008). An associated problem is thus the fact that although the basic definition of field does not preclude its use at a variety of levels, many of the conditions allowing for relative autonomy of a given field are themselves conditional upon the existence of a state able to distribute credentials and regulate entry to certain professions or circles (Bigo and Madsen, 2011; Vauchez, 2011).

Indeed, the relative autonomy of transnational fields vis-à-vis national fields remains one of the key theoretical challenges for transnational field theory – particularly for those works examining fields that emerge as a result of the intersection of national fields. Remarkably, some of the early importers of Bourdieu's ideas into the IR discipline remain skeptical or at least cautious as to the very existence of *autonomous* transnational fields (Cohen, 2018). This is the case with Dezalay and Garth (2002) and, later on, Vauchez (2011), who have argued that international and transnational interactions remain highly anchored to domestic struggles, and that the autonomy of most fields is highly dependent on existence of some form of statehood. Vauchez (2011) thus notes that even highly institutionalized international polities and settings (for instance, the EU) are not fully governed by a self-referential logic – and that most of their professionals continue to be embedded in national fields of power. The author notes that transnational fields can often be conceptualized as "weak" fields, displaying many of the characteristics of emerging fields – namely a limited autonomy vis-à-vis "neighboring fields" (including, but not only, national fields), as well as a less obvious or institutionalized structure. Similarly, the work of Dezalay and Garth, with a focus on legal practices, has tended to emphasize that many actors use transnational forums to affect their positions in the national arena (and vice versa), concluding that international fields generally remain in an embryonic state (Dezalay and Garth, 1996; 2002). Conversely, other authors are more confident in

the (theoretical) possibility that transnational fields can eventually become autonomous spaces. This is, for instance, the case with the line of inquiry initiated by Bigo (2005), which focuses on the emergence of a security field cutting across organizational but also national boundaries, thus involving a wide range of heterogeneous actors.

Echoing the lack of consensus regarding such questions, we find that, as applied by IR scholars, the notion of field comes in many guises, giving rise to a variety of interpretations. For the purposes of this dissertation, the notion of *fields of transnational governance*[64] as proposed by Sending (2015) features as a particularly apposite application of the concept. The construct is developed by the author as a heuristic tool to theorize global governance,[65] by bringing to the fore the dynamic and interactive nature of governance arrangements, as well as by casting light on recognition dynamics – understood as "the mechanisms through which relations of authority are established and perpetuated" (Sending, 2015, p. 24).

According to this approach, governance fields are those organized around governance objects – that is, internationalized and institutionalized themes (such as peacebuilding, development and health etc.). Transnational fields of governance encompass the totality of actors somehow engaged in the definition, construction and management of governance objects – including international organizations but also states, epistemic communities, NGOs and advocacy networks etc. Fields are thus populated by a constellation of collective but also individual actors brought together by processes of transnational governance. The construct is therefore uniquely posited to capture the multilevel and multi-actor character of transnational governance, as well as to make sense of the emergence and structuration of transnational social spaces.

Thus, field theory essentially represents a mode of analysis – that is, a strategy to articulate and construct a research object and uncover dynamics that would otherwise remain invisible. Applying a field perspective to the study of a social space requires identification of an arena of struggle (i.e., a set of stakes over which a range of actors compete) and then its topography – that is, mapping out its structure, identifying a hierarchy of positions, and determining its boundaries and relative autonomy (Pouliot and Mérand, 2013; Sending, 2015).

Engaging on a topographic exercise of the contours and structure of the education-for-development field is well beyond the scope and objectives of this dissertation. However,

---

[64] Note that the notion bears a resemblance to the notion of *transnational governance fields* as advanced by Djelic and Sahlin-Andersson (2006). However, the theoretically eclectic approach orienting Djelic and Sahlin-Andersson's (2006) work (as opposed to Sending's (2015) heavy reliance on Bourdieu's field theory) renders the two approaches incomparable.

[65] Understood, in turn, as a form of soft regulation resting upon non-binding rules – that is, in which compliance is driven by standards, guidelines, normative pressures, regulatory competition or socialization dynamics, rather than coercive pressures.

the tenets and insights advanced by field theory represent a useful departure point to one of the potential limitations of the literature reviewed in the prior two sections – namely, the limited emphasis placed on the interaction of IOs within processes of norm emergence and indicator-making. While literature on *internal* normative change, and on the quantification exercises of *specific* IOs, has proliferated over the last years, it remains less clear how different organizations sharing a mandate on a given issue-area negotiate and build consensus in norm- and indicator-building processes. A field approach brings to the fore the question of inter-organization interplay and forces us to think in terms of organizational interdependence. For the study of the making and the monitoring of SDG4, this represents a particularly apposite approach.

**Wrapping up**

This chapter has brought together three distinct areas of scholarship (international norm-building processes; the construction and impact of global indicators; and the relational character of global governance) in order to articulate theoretically the interface between global goals and international organizations as an object of study. It has thus identified two complementary *angles* through which to investigate global goal-setting practices. On the one hand, the chapter has firstly argued that global goals can be productively analyzed as instances of international norm emergence, and that they are particularly well suited to unearthing the role of IOs in such processes. On the other hand, and given the particularities of the SDGs, it has also made the case for an understanding of SDG4 as a major quantification exercise with the potential to substantially affect the ecology of organizations engaged in promulgation and production of education indicators but also of education datasets. Finally, the chapter has introduced field theory (as advanced by the Bourdieusian tradition) as an appropriate *lens* to incorporate a relational perspective, as well as to capture the multi-actor character of transnational governance processes characterized by fuzzy hierarchy lines – as has been the case with negotiation and monitoring of SDG4.

# Chapter 4

## Research design and methods

**Introduction**

This dissertation aims to examine the negotiation and monitoring of Sustainable Development Goal 4 by bringing to the center of the analysis the culture and interests of the international organizations participating in these endeavors, as well as the institutional architecture in which inter-organizational interactions took place. This chapter discusses the empirical strategy devised to address this objective, in an attempt to make the methodological and analytical decisions underpinning this research transparent and explicit.

The chapter is structured into two main sections. This first section of the chapter discusses the research strategy employed to generate the corpus of data of the dissertation – including the strategies used to select and recruit these informants and the rationale behind the design of the interview guides. The section also reflects on the specificities of on-line interviewing and the methodological implications and challenges entailed by this medium. Finally, the section discusses two additional sources of information that were used to triangulate the insight gained from the interviews – namely, documentary data and non-participant observation.

The second section of the chapter presents the analytical strategy used to make sense of the collected data – a case-centric variant of process tracing. It therefore provides an overview of the main tenets of this analytical approach, as well as of its adaptation for the purposes of interview-coding. The section also includes a brief comment on the use of thick description as the strategy employed to present the main findings of the research.

## 4.1. Overview of the data-generation process

### 4.1.1. "Studying up": introductory remarks on the practice of elite interviewing

The research relies on semi-structured interviews as the main source of data. Over the period 2016 to 2020, 99 semi-structured interviews were conducted with a wide range of stakeholders who had been or remained involved in the negotiation of the SDG4 agenda and the production of associated metrics. Typically, the targeted interviewees were bureaucrats and negotiators occupying positions of responsibility or authority within their own organizations. Thus, most of the interviews conducted in the context of this research fell within the *elite interviewing* category. The bulk of the interviewees can indeed be characterized as *elite education policy actors* in the sense advanced by Lingard, Sellar and Baroutsis (2015) – a mix of global policy-makers and global policy-technicians with

"cosmopolitan sensibilities that prioritize outwardly-oriented global perspectives as a frame for understanding the national/local" (p. 26) and which exhibit the specific capitals (symbolic, social, cultural) associated with policy-making in transnational spaces.

The very notion of elite interviewing is certainly a highly contested one. The dichotomy between elites and non-elites has indeed been called into question (Smith, 2006). As observed by Harvey (2011), "the term elite can mean many things in different contexts, which explains the range of definitions" (p. 433). Thus, some authors rely on criteria relative to the power or influence of a given individual within the business and political apparatus of a given country, whereas others emphasize specialized knowledge and professional credentials. Finally, some authors rely on relational criteria – emphasizing the social distance between the researcher and the interviewer. In this sense, it is possible to argue that the specificities of elite interviewing do not lie so much in the attributes of the interviewees, but in the challenges posed by the interpretation of the resulting data. This line of reasoning is advanced in Dexter's classical text on elite and specialized interviewing, where he notes:

> In the standardized interview, the typical survey, a deviation is ordinarily handled statistically; but in an elite interview, an exception, a deviation, an unusual interpretation may suggest a revision, a reinterpretation, an extension, a new approach. In an elite interview it cannot at all be assumed – as it is in the typical survey – that persons or categories of persons are equally important. (1970, p. 25).

The subjects interviewed in the context of this dissertation conform to this broad understanding of elites. The decision-making power and authority exhibited by many of the interviewees has been an issue with key implications on access strategies, the design of the interview protocol, and data analysis. As noted by different authors, the practice of elite interviewing poses a number of challenges relative to the access, trust-building, tone-gauging and interpretation that require further research (Aberback and Rochman, 2002; Harvey, 2011; Morris, 2009). In the light of this, I refer in this section to the strategies that I have adopted to address such challenges.

*4.1.2. Constructing the pool of interviewees*

In line with the process-tracing approach guiding the research (cf. *Overview of the analytical strategy* below), the selection of interviewees has relied on a non-probable sampling approach. As argued by Tansey (2007) in his discussion of the use of elite interviewing, this is the most appropriate approach when the objective of the study is to reconstruct an event (or set of events) rather than to generalize or draw inferences about a wider population. Since the objective of this dissertation is precisely to get a fine-grained understanding of the negotiation of the SDG4 targets and the associated metrics and reporting protocols, the main concern driving the selection of interviewees was not representativeness, but rather to obtain the perspective of the totality of actors with a significant involvement in such a process.

However, the porous and unbounded nature of the SDG4 case study rendered the identification of relevant actors a challenging exercise. This was the case for two different, albeit interconnected, reasons. First, the range of organizations that provided input to the post-2015 education debate and the construction of SDG4 targets and indicators was notoriously vast. Virtually all the multilateral and bilateral development partners and international organizations with a mandate in education had been formally involved in the consultation and deliberation *fora* set in place for the purpose of negotiating a new agenda and defining the associated metrics. Similarly, a wide array of think tanks, NGOs, CSOs, experts and private sector representatives had also taken part in such deliberations. However, the intensity and duration of the engagement of such actors was distinctly heterogeneous. Some organizations had provided their input on punctual occasions, whereas others had been more consistently present in the myriad of deliberation venues and initiatives set up. In addition, not all the organizations *formally* included in these deliberation exercises had taken a genuine interest in the process – there was considerable variability in the relative centrality of SDG4-related debates within the organizational agendas of the participating IOs.

A second difficulty emerged from the fact that identifying the *individual* negotiators and representatives involved in the SDG4 negotiation process was not always a straightforward task. While some meetings, working groups and consultations offered detailed lists of individual participants, this was not always the case. In addition, over the course of the first interviews, it became apparent that the knowledge and familiarity with the post-2015 process exhibited by the different staffers of a given organization was highly unequal. The largest organizations had designated one or a few staffers responsible for following the process, providing input and briefing their colleagues and partners on the progress made by the negotiations. However, identifying such people was not always an easy endeavor – for such "SDG4 specialists" were rarely the only negotiators designated to participate in the meetings, working and drafting groups, committees and consultations. Some initial interviewees suggested that some of the most knowledgeable actors in certain organizations were not necessarily those attending relevant meetings on a regular basis, or formally featuring in committees and working groups. There was therefore a high risk of omitting relevant figures.

It is true that such difficulties were much more pronounced in relation to the negotiation of the SDG4 targets than was the case for the indicator-production efforts. On account of its (allegedly) technical nature, the latter concerned a more limited group of organizations and experts, and the debates and negotiations took place in a more limited number of *fora* and initiatives. It was thus possible to identify a number of "core" actors (assessment experts, representatives of statistical agencies, IO staffers trained in statistics, and so on) that participated in the range of meetings and took part in the succession of working groups with a role in the conceptualization and production of the SDG4 indicator framework. The quantification debate appears to exhibit better-defined (or more formalized) contours – a feature that rendered the selection of interviewees who were most likely to have an in-depth knowledge of the intricacies of the process.

In light of this, in order to identify potential interviewees I relied upon two different non-probability strategies – corresponding to the two research objectives guiding the thesis. In relation to the *first objective*, and in order to reconstruct the negotiation of the SDG4 targets, I relied primarily on a snowballing sampling strategy. As noted by Tansey (2007), while this sampling strategy is typically associated with the study of "hidden populations" (e.g., those engaged in criminal activity and, consequently, not fully visible), it is also an appropriate research study for the study of elite populations, in those contexts "where the most influential political actors are not always those whose identities are publicly known" (p. 770). Thus, I started the snowballing process from three separate departure points – namely, the List of Members of the EFA Steering Committee as of 2015, the List of Participants in the World Education Forum 2015 (specifically, the UN agencies, NGO and Private sector categories), and the list of Members and Contributors of the Youth, Education and Culture thematic cluster of the Open Working Group. In the case of the WEF 2015 register of participants, given its extension, the list was further parsed in order to identify those members whose engagement with the post-2015 process appeared to be minimally sustained over time, and to exclude those individuals whose participance appeared to be honorary (e.g., senior figures in IOs not directly concerned with education debates). In the case of the Youth, Education and Culture thematic cluster, the list was narrowed down to those representatives of organizations with an agenda on education. This gave rise to a first list of prospective interviewees. A final screening was conducted in order to exclude those members whose contact details were not publicly available.

The prospective interviewees identified through this first round were approached by an introductory email explaining the rationale of the interview. I attached a letter to this email describing the objectives of the research and giving details of the data-handling procedures, which ensured the confidentiality and anonymity of the information provided by the participants. This is consistent with the guidance on elite interviewing provided by the Economic and Social Research Council (ESRC) *Framework for research ethics*[66], which notes:

> In elite interviews it is often argued that formal written consent is not necessary because by consenting to see the researcher, the participant is in fact giving consent. However, all such participants should receive an initial letter giving the name and status of the researcher carrying out the study, a brief rationale of the study including its purpose and value and why the individual is being invited to take part. The person interviewed should be aware of what will happen to any findings, whether the data will be shared with others, and whether he/she will be identified. (ESRC, 2015, p. 42).

The initial sample of respondents that accepted the request became the basis upon which I relied to identify further potential interviewees. Thus, during the early stages of the research, I routinely asked my interviewees for recommendations and interview leads – individuals they identified as particularly knowledgeable about the questions discussed

---

[66] I relied on the ethical standards provided by the ESRC due to its strong reputation and because the ethical guidance offered by UAB did not discuss the specificities of elite interviewing.

over the course of our conversation. The snowballing strategy made it possible to expand the pool of interviewees. Additionally, being able to mention prior interviewees proved helpful in reducing the ratio of interview declines. The number of interviews declined was particularly high during the first stages of the research, when interviewees were approached through "cold calls". At the same time, and in order to avoid the risk of selection bias associated with snowballing (cf. Martin, 2013), I prioritized those (suggested) interviewees whose position was not *structurally equivalent* to those already interviewed (e.g., those occupying a similar rank and position within the same organization). In addition, I proactively asked about specific organizations which were underrepresented in the emerging corpus of interviewees. This is in line with the mitigation strategies advanced by Ritchie, Lewis and Elam (2003) and Bleich and Pekkanen (2013) and prevented the over-representation of those constituencies that were easier to access or particularly eager to be interviewed (for instance, as a result of their discontent with the final outcome of the agenda).

In relation to the *second research objective*, relative to the production of globally-comparable learning data, I relied on a purposive sampling strategy (also known as judgement sampling). Purposive sampling represents a criterion-based approach to the selection of interviewees, in which participants are selected "because they have particular features or characteristics which will enable detailed exploration and understanding of the central themes and puzzles which the researcher wishes to study" (Ritchie et al., 2003, p. 79). This strategy is appropriate to ensure that all the individuals whose perspective and experience of the process is judged to be *unique* and *relevant* are included in the sample. The approach is therefore strongly based on the researcher's criteria to identify those individuals more likely to provide insight into the research puzzle (Mason, 2002; Tansey, 2007). It is not oriented towards ensuring the representativeness of different groups, rather it aims to capture the perspective of all those individuals displaying a feature that the researcher judges to be of interest (e.g., affiliation, specific position within the organization chart, singular experiences, etc.).

Thus, in order to gain insight into the debates relative to the production of cross-national learning data, I reviewed the different experiences set up by the UIS since the mid-2000s. I also reviewed the technical reports and the documentation produced in the context of the Learning Metrics Task Force, the Learning Metrics Partnership and the Global Alliance to Monitor Learning and the Technical Cooperation Group in order to identify those organizations and individuals that appeared to have been more consistently involved in such debates – either through their participation in key meetings or as authors of reports or background documentation. This group of experts included a number of UIS staffers, assessment and learning experts employed by other IOs and donors, representatives of those organizations in charge of large-scale learning assessments, think tank and country representatives. Given that the quantification efforts have been led by a relatively reduced group of experts, and since public documentation on the process was distinctly complete and transparent, the identification of a core group of key players was more feasible compared to the first research objective. Consequently, I did not resort to

a snowball strategy in order to identify potential interviewees (even if I took advantage of name-dropping and spontaneous recommendations) – rather, I defined the sample frame on the basis of theoretical considerations and a prospective analysis of relevant documentation. Accordingly, and in contrast to the first research objective, on this occasion I prioritized positional criteria (relative to the position, experience and occupation of the potential interviewees) rather than reputation criteria (relative to the presumed relevance of actors according to their peers) (cf. Tansey, 2007; and Farquharson, 2005, for a discussion of such categories). The prospective interviewees identified through this strategy were also approached via email and provided with an information letter discussing the purpose of the interview and the management of anonymity and confidentiality.

The combination of these methods gave rise to a final corpus of 99 interviews, with an average length of one hour. Table 4.1 below offers a breakdown of the pool of interviewees according to six basic categories – namely:

- UN agencies (UN[67]) – senior and junior officials and analysts employed by UNESCO (including IIEP-UNESCO and the UIS), UNICEF, UN Women, UNHCR, etc.

- Large-scale assessments (LASS) – representatives and experts from organizations in charge of large-scale assessments (in turn, a combination of IOs[68], regional networks, NGOs and expert consortia).

- Multilateral and bilateral donors (DON) – including staffers and officers from the World Bank, the GPE, and bilateral aid agencies.

- Civil society organizations and non-governmental organizations (CSNGO) – organizers and negotiators representing these two constituencies.

- Country representatives (CR) – senior civil servants and government officers in representation of Member States.

- Private sector and experts (PRI) – representatives from the private sector, philanthropic and third-sector organizations and foundations, think tanks and research institutes formally involved in the post-2015 debate or the production of learning indicators.

---

[67] Note that the bracketed acronyms correspond to the nomenclature used to quote interviewees in Chapters 5 and 6.

[68] Note that representatives of IOs and NGOs selected on account of their role in the production of learning data are included in this category (LASS) rather than in the UN and CSNGO categories.

**Table 4.1. Pool of interviews – disaggregated by categories**

| Interview group | Total | Research objective 1 | Research objective 2 |
|---|---|---|---|
| UN | 33 | 21 | 15 |
| LASS | 16 | 3 | 16 |
| DON | 14 | 7 | 12 |
| CSNGO | 16 | 11 | 5 |
| CR | 6 | 3 | 3 |
| PRI | 14 | 7 | 8 |
| *Total* | *99* | *52* | *59* |

Source: Author's elaboration.

*Note:* Please note that the second column does not (necessarily) correspond to the arithmetic sum of the third and fourth columns – given the fact that some interviewees provided input to both research objectives.

Finally, it should be noted that, as a result of the non-probable sampling approach used in the study, determining the saturation point required an element of judgement. According to the specialized nature regarding the use of interviews for the investigation of political phenomena, saturation is reached when respondents refer to the same causal processes as former interviewees, when there is agreement across networks, or when disagreement becomes predictable – i.e., when no new analytical insight emerges (Bleich and Pekkannen, 2013). However, such a saturation point is not incompatible with the persistence of uncertainty over specific events or points of the discussion. In the context of my research, such uncertainties stemmed from the fact that certain interviewees (particularly, top-level bureaucrats) remained inaccessible (e.g., if contact data were not available, or if interview requests were declined or ignored). In order to compensate for such gaps, follow-up interviews were carried out with seven respondents from the first rounds of data collection, conducted mostly over the last year of data collection (2019/2020). These informants were selected on the basis of the knowledge and openness they had exhibited in the first interview, as well as on the basis of their role in some of the events (meetings, preparation of reports, and so on) identified as key on the basis of interview data. In the context of these follow-up interviews, I requested clarification, elaboration or reaction to my emerging insight and efforts to "fill" the research gaps relative to the lack of access to certain interviewees. Additionally, these follow-up interviews were also instrumental in gaining up-dated approaches to events and developments that had occurred over the course of the research.

*4.1.3. The interview guides*

The interviews relied on two main interview guides, corresponding to the two main research objectives, that were developed and refined over the course of the research. Both questionnaires can be consulted in the Annex of this dissertation. It should however be

noted that such interview questionnaires were used more as a guide than an interview protocol, for they had to be tailored to the specificities of each interviewee – that is, adapted to the degree of familiarity with the process, the specific venues they had participated in, the nature of their engagement, and so on. As noted in the previous section, the experience and position of the interviewed individual was typically idiosyncratic. Thus, the individuals had been selected given their role in critical junctures of the process, because they were recognized as prominent experts in certain areas, or since they were among the few staffers with in-depth knowledge of the post-2015 process or of learning metrics within a given organization. This rendered impossible the use of a strict protocol interview – the interview guides were employed essentially as an extensive battery of open-ended questions oriented towards preventing overlooking relevant questions. This is what Merriam and Tisdell (2016) characterize as the *semi-structured format*.

The interview guide, relative to the first research objective (see IG1 in the Annex), was organized along six main sections or modules. These corresponded broadly to the themes addressed by the specific objectives associated with the research goal, and most of them contained a combination of *knowledge questions* (i.e., questions oriented towards eliciting the respondent's factual knowledge of the process), *experience and behavior questions* (oriented towards understanding the role of the interviewee in relation to the discussed events) (cf. Patton, 2015).

- The first module (*Background questions*) was oriented towards gaining insight into the professional trajectory of the interviewee and the responsibilities of their position. In addition, it aimed to identify the ways in which the respondent had been engaged in the post-2015 debate, and to determine the intensity of such engagement and the interviewee's knowledge and familiarity with the different strands of the debate, in order to tailor the questions accordingly.

- The second module (*Organizational engagement*) was oriented towards determining the different deliberations, consultations or debates to which a specific organization had contributed over the course of the post-2015 debate, the motivation behind such engagement, and the relative priority of the post-2015 debate within the organizational agenda.

- The third module (*Organizational priorities*) was oriented towards discerning the priorities and preferences of the organization in relation to the *content* of the new agenda (i.e., normative preferences), as well as to understanding the extent to which different organizations perceive the final SDG4 framework as aligned to their own agendas and priorities.

  In combination with the second module, such questions responded to the specific objectives 1.2. and 1.4. – i.e., reconstructing the rationales and priorities driving the participation of different organizations in the process and examining the main focus of debate, contention and (open or covert) disagreement.

- The fourth module (*Overview of the process*) was oriented towards gaining insight into the sequence of events that had shaped the final outcome (relevant episodes, points of tension, etc.), as well as to discerning the priorities of the organization in relation to the *architecture* of the process.

- The fifth module (*Partners and network*) was oriented towards determining patterns of inter-organizational collaboration, and to identifying those actors more generally perceived as influential.

  In combination with the fourth module, such questions responded to the <u>specific objectives 1.1., and 1.3</u> – i.e., reconstructing and analyzing the sequence of inter-governmental, and intra- and inter-organization negotiations that shaped the elaboration of the final list of targets and indicators, and identifying the varying degrees of influence exerted by different organizations, as well as the key *locus* of decision-making.

- The sixth module (*Closing questions*) was oriented towards identifying other potentially relevant interviewees (as requested by the snowballing sampling approach) and wrapping up the interview.

The second interview guide, relative to the production of learning metrics (see IG2 in the Annex) was similarly organized along the specific research objectives. However, unlike IG1, this second guide contained some modules targeted exclusively at some interviewees. The list of modules is as follows:

- A first module (*Background data and individual engagement*) oriented both towards gaining insight into the area of specialization of the interviewee and discerning their familiarity and knowledge of a number of specific initiatives judged of relevance.

- A second module (*A renewed emphasis on learning assessments*) oriented towards identifying the interviewee's perceptions of the opportunities, challenges and risks posed by the growing prominence of learning metrics within the education-for-development realm, as well as the changes that such a shift had entailed in the agenda and organizational priorities of different organizations.

- A third module (*The role of the UIS*) oriented towards reconstructing the impact of the aforementioned changes in the authority, visibility and internal organization of the UIS – as the custodian agency of most of the learning-related indicators.

  In combination with the second module, such questions responded to the <u>specific objectives 2.1. and 2.2.</u> – i.e., to analyze the ways in which the impetus given to learning metrics in the SDG4 agenda has penetrated into the institutional agendas and priorities of the main organizations with a role or mandate in the production and harmonization of assessment data, and to examine how the commensuration and coordination needs brought about by the new monitoring mechanisms have affected the relations of interdependence among the producers and the harmonizers of learning data.

- A fourth module (*The work of GAML*) oriented towards understanding the technical and political challenges posed by the production of learning-related indicators. Such questions responded to the specific <u>objectives 2.3 and 2.4.</u> – i.e., to examine how the new quantification needs have transformed the hierarchization dynamics governing the assessment field, and to identify the strategies mobilized by these organizations in order to preserve or access dominant or influential positions.

- A fifth module (*Closing questions*) was oriented towards wrapping up the interview and identifying potential documentation of interest.


*4.1.4. Conducting research at a distance*

Given the transnational nature of the field that constitutes the object of this research, the professionals and bureaucrats that populate it are geographically scattered. Not only they are employed in institutions hosted by a wide range of cities (Paris, New York, Washington DC, Brussels, Montreal, etc.), but in addition, most of the interviewees travel frequently on account of their professional responsibilities. In light of this, and for feasibility reasons, the bulk of the interviews were conducted online – on the basis of videoconferencing software (with the exception of 10 face-to-face interviews and two interviews conducted over the telephone).

Over the last decade, the specificities of online interviewing have become the object of methodological reflection, with the publication of some volumes providing guidance on such methods and reflecting on the effects of the medium of data collection (e.g., O'Connor and Madge, 2017; Salmons, 2015). The potential of this modality is generally acknowledged – for it represents a way of overcoming financial and logistical considerations that have typically exerted a constraining effect in the selection of interviewees, while preserving the benefits of visual of cues and synchronicity (as opposed to the use of telephone interviews or interviews over email). However, it has also been noted that the use of web conferencing is not without problems – relative to ethical issues (for the reliance on commercial software might compromise confidentiality), the risk of logistical complications (e.g., the risk of computer breakdown, poor connection, etc.), or the increased difficulty to build rapport between the interviewer and the interviewee – as opposed to face-to-face interviews (cf. Deakin and Wakefield, 2014; Merriam and Tisdell, 2016; or Sedgwick and Spiers, 2009). Some authors have thus concluded that more research is needed in order to understand the pluses and minuses of computer-mediated interviewing, as well as the most fruitful strategies to maximize the quality and value of such exchanges (Salmons, 2015).

Over the course of my research, I had the opportunity to reflect on the opportunities and risks associated with online interviewing, and on potential strategies to mitigate the limitations imposed by this medium. On the plus side, online interviewing appears to have some advantages over face-to-face (or telephone) interviews that are well beyond

logistical and financial considerations. These advantages relate first to the ease of recruitment – the possibility of conducting the interview online helped some interviewees perceive the appointment as less of an "interruption" in their schedules and calendars – as they could simply pencil in a tentative date and arrange for the specific time at the very last moment. Secondly, computed-mediated conversations allowed interviewees to check their notes, calendars, webpages, pieces of research or contact lists. Considering the retrospective nature of the interviews, the possibility to consult documentation helped to increase the "confidence" or ease of the interviewees, and notably improved the richness and detail of the elicited information.

At the same time, the reliance on videoconferencing software to conduct interviews also appeared to have important downsides. In line with the insight offered by specialized literature, most of the difficulties to build rapport stem in fact from technological limitations. Poor or unstable bandwidth typically results in audio and video "freezes", which greatly limited the fluidity of the conversation. On some occasions, the poor quality of the audio created inter-comprehension problems[69] and rendered the management of pauses, interruptions and speaking turns particularly challenging. Such difficulties could generally be remedied by turning off the cameras – even if this resulted in a loss of nonverbal cues. An additional difficulty stemmed from the loss of "off-the record" time. While in face-to-face interviews the recording time is clear to the interviewee (who can see how the interviewer manipulates the digital recorder), this is less obvious in the context of an online interview. My experience with face-to-face interviews (in this and other studies) suggests that certain interviewees (those in junior positions, those who are particularly tense over the possibility of being identified) use the off-the-record time to clarify certain questions, add nuance or reveal pieces of information that they judge too delicate or too "personalistic". In order to compensate for the loss of these moments, I proactively offered the interviewees the possibility to turn off the recorder at any moment they chose. A final challenge posed by the use of videoconferencing was the frequency of no-shows or last-minute cancelations. It is difficult to ascertain to what extent such absentees can be directly attributed to the medium – however, some accounts suggest that absence of a physical meeting point tended to facilitate no-shows (Deakin and Wakefield, 2014). In order to mitigate such difficulties, I resorted to a relatively intensive use of reminders and, where possible, took the liberty to create a calendar event linked to the invite.

### 4.1.5. Complementary sources of information

In their seminal book on the foundations of process tracing, Beach and Pedersen (2013) argue that the validity of the insight gained by such methods depends greatly on the quality and accuracy of the collected data (*observations* according to the terminology

---

[69] This might have been compounded by the fact that, frequently, the language of the interview (English) was a second language for both the interviewer and the interviewee.

used by the authors). The authors do in fact issue some cautionary notes regarding the use of elite interviews, noting that "it is vital to consider who is being interviewed and his or her motivations for doing so" (p. 126). The authors go on to outline a number of risks relative to this specific source of evidence – most notably, the fact that respondents are likely to exaggerate their own centrality in a given political process or to offer self-serving accounts, the lack of reliability of interviewees relying on hearsay and secondary sources, and finally, the temporal gap between an event and the moment of the interview.

> Participants interviewed immediately after a negotiation can be expected to have a reasonable recollection of the details about what took place, but lapses of memory over time will result in observations that are less reliable. More insidious is the risk that participants will, by reading other accounts and talking with other participants, change their interpretation of what happened to match other accounts, resulting in potential bias in our observations if we interview these participants long after the fact. (2013, p. 135).

The *temporality* factor was challenging in the context of my own research. Most of the fieldwork relied on retrospective accounts – thus requiring informants to reminisce about past events, some of which dated back to almost five years. Process tracing and thick description benefit from nuanced and detailed-oriented explanations (rather than impressionistic accounts) – however, eliciting such information in relation to remote events proved a highly challenging endeavor. The ability of my interviewees to dwell on the past and elaborate on the process differed greatly and was rather unpredictable. Their recall depended on a wide number of variables, including the intensity and extension of their involvement within the post-2015 process, and their professional trajectory and personal connections, seniority, individual dispositions and interest in such themes, among other factors. Sharing the script ahead of the interview ended up proving to be a useful strategy to overcome this challenge. While some of the literature on elite interviews explicitly advises against such practice, in the context of my research this turned out to be a helpful approach since, by giving interviewees the opportunity to think about their responses in advance, it facilitated the production or rich and particularized accounts (and avoided the tense atmosphere that is likely to emerge if the interviewee feels unprepared). At the same time, script-sharing was not without costs – it sometimes led to stilted accounts while reducing the spontaneity that might characterize a less intensely scripted conversation.

In order to compensate for such limitations, I turned to a triangulation strategy – as advised by Beach and Pedersen (2013) and other authors focusing on the challenges of interviewing for the study of political phenomena or "studying up" (i.e., researching elites) (see Bleich and Pekkanen, 2013; or Nader 1972). Therefore, and in addition to the follow-up interviews discussed above, two complementary sources of data were used to double-check, add nuance, clarify and expand the insight gained in the context of the interviews. First, I reviewed a broad *corpus* of documentation collected over the course

in an iterative way (n=130)[70]. The review of such documentation was crucial in order to corroborate the input offered by interviewees, as well as to add nuance and clarity to their recollections.

This *corpus* of documents included official reports, memos and minutes produced by the major education agencies, position papers, research reports and prospective studies produced or commissioned by the organizations participating in the debate, outcome documents prepared in the context of consultations, political *communiqués*, public statements made by different stakeholders, blog posts published by negotiators, meeting agendas and background documents, policy briefs, evaluation reports and meeting presentations, *i.a.* The collection of such documents had an iterative nature and did not follow a pre-determined protocol or search strategy – with the exception of an initial round of online searches which was conducted in 2015, with the objective to collect and systematize all the relevant documentation produced in the different strands of the post-2015 debate. To this end, I scanned the UNESCO repository (UNESDOC Digital Library) searching for specific keywords, the official web pages of the organizations formally involved in the post-EFA debate, and the Sustainable Development Knowledge Platform (specifically, the section on the post-2015 process, which operates as a clearing house of the documentation produced in the context of the OWG and the Intergovernmental negotiations). From there, I conducted regular screening of the web pages mentioned above along with searches in other institutional portals (e.g., the UIS web pages dedicated to Global Alliance to Monitor Learning (GAML) and the Technical Cooperation Group (TCG), or the World Education Blog curated by the GEMR). A number of ad hoc searches were also conducted in response to new developments of the debate and to imprecisions in the accounts offered by the interviewees. These searches were also to test and substantiate the emerging analytical insight.

To complement this I also relied on observation data in order to contextualize and decode the insights gained on the basis of interview and documentary data. Specifically, I attended five multi-day events directly related to the monitoring and follow-up of the SDG4 agenda, and the production of its associated metrics – namely:

- *2nd Meeting of the Technical Cooperation Group on the Indicators for SDG4-Education 2030* – 16 to 18 October 2016, Madrid.

- *3rd Meeting of the Global Alliance to Monitor Learning* – 11 and 12 May 2017, Mexico City.

- *5th Meeting of the Global Alliance to Monitor Learning* – 17 and 18 October 2018, Hamburg.

- *Global Education Meeting 2018* – 3 to 5 December 2018, Brussels.

---

[70] The documents included in this *corpus* are marked with an asterisk in the *List of references* of this dissertation.

- *6<sup>th</sup> Meeting of the SDG-Education 2030 Steering Committee* – 11 and 12 March 2019, Paris.

I attended these events as a non-participant observer. Such observations were not designed as an ethnographic endeavor, but as a means to gain familiarity with the community that constituted my focus of study. The observations allowed me to gain insight into the inner logics and rhythm of global gatherings, and to understand the "on-the-ground" meaning and the significance that my interviewees attributed to certain objects – e.g., the Steering Committee, the provisional list of indicators or the Framework for Action. The insight gained from such encounters played a crucial role in the subsequent analysis and de-codification of the interview data, in that it enabled me to interpret the recollections of my interlocutors in a more context-sensitive way – one that took into account their worldviews and the meanings they attributed to the object of this research.

Overall, my approach was inspired by the principles and propositions of what Schatz (2009) has characterized as an *ethnographic sensibility* to the study of political phenomena. As advanced by Schatz (2009), ethnography can be understood as a ground-level method (typically equated to participant observation) but also as a particular sensibility that focuses on insider perspectives. An ethnographic sensibility is thus to be applied not only to observation practices, but also to the analysis of other data sources – as a means to elucidate and make sense of the logic, motivations and beliefs guiding decision-makers:

> [*T*]he term sensibility goes at least partway to transcending artificial distinctions between fieldwork and deskwork, between research site and site of analysis, between researcher and researched, and so on, distinctions that are hard to sustain in a world that defies these binary distinctions. It also avoids reducing ethnography to the process of on-site data collection. Sensibility implies epistemological commitments that are about more than particular methods. In this sense, an ethnographic study usually employs multiple tools of inquiry. (Schatz, 2009, p. 6).

In the context of these field observations, I took extensive notes not only relative to the "central stage" (that is, the formal deliberations and exchanges, as scheduled by the meetings' agendas) but also to hallway exchanges, informal chats and side conversations held over recesses and breaks. While, given its exploratory and a-systematic nature, observation data has not been used to construct fieldwork vignettes, such notes proved useful both to prepare for and to interpret the interviews – that is, in order to interrogate data in a way sensible to the views and universes of meaning of those inhabiting the education-for-development field.

My participation in such events was also instrumental in expanding the pool of interviewees and in building a solid rapport with my respondents. Thus, the events I attended proved a useful opportunity to introduce myself to prospective interviewees – a move that hypothetically improved the rate of acceptance of interview requests.

Additionally, being perceived as someone with a degree of familiarity with the SDG4 architecture rendered it easier to build a productive rapport with my interviewees. The process of trust building also benefitted from my sporadic work as a UNESCO consultant during the 2019-2020 period[71]. Thus, whilst the atmosphere of my first interviews was frequently stilted (and my limited knowledge of the codes, assumptions and "common knowledge" of the international bureaucracies was often an obstacle to the open flow of information), my engagement with transnational *fora* contributed to create a more productive interview situation.

## 4.2. Overview of the analytical strategy

### 4.2.1. A process tracing approach

The analytical strategy employed to make sense of the data was informed by the principles of *process tracing,* understood as a *within-case method of analysis*[72]. According to George and Bennett's (2005) oft-cited definition, process tracing can be defined as an analytical strategy attempting "to identify the intervening causal process – the causal chain and causal mechanism – between an independent variable (or variables) and the outcome of the dependent variable" (p. 206). As argued by Waldner (2012), there is, in practice, limited agreement on the very definition and limits of this approach (to the point that it has become something of an "umbrella term" encompassing a multiplicity of strategies). However, some of the most distinguishing features of process tracing appears to be its focus on *sequences of events* (i.e., concatenation) and a *mechanism-based* understanding of causation:

> The core idea of process tracing is concatenation. Concatenation is the state of being linked together, as in a chain or linked series […] In process tracing, one concatenates causally relevant events by enumerating the events constituting a process, identifying the underlying causal mechanisms generating those events, and hence linking constituent

---

[71] From January 2019 to June 2020, I conducted consultancy work for the Education Research and Foresight team of UNESCO's Education Sector on an intermittent basis.

[72] While, in the context of this dissertation, process tracing is approached as an *analytical strategy*, it should be noted that process tracing is sometimes characterized as a *research method* or *methodology*. I refrain from doing so in light of Colin Hay's (2016) reflections on the difficulty of translating the principles of process tracing into specific methodological strategies. In a recent essay, Hay has taken issue with "the gulf between the ambition and expectation of process tracing […] and the methodological means to deliver that ambition on the other hand" (p. 502). The author argues thus that it remains unclear whether process tracing is associated with some form of methodological specificity and contends that, while self-declared process-tracers privilege certain data-gathering and inference strategies (Bayesian logics, diachronic analysis, etc.), there is no a priori reason why process tracing needs to rely on these methods. According to the author, the ambition to gain insight into the causal mechanism and processes at play behind a given outcome, can be satisfied by recourse to a wide array of methodological options. Hay's plea for methodological pluralism is synthesized in his observation that "process-tracers […] do not have a monopoly over process tracing" (p. 501).

events into a robust causal chain that connects one or more independent variables to the outcome in question. (Waldner, 2012, p. 69).

A similar argument is advanced by Bennett and Checkel, who have recently revisited and refined George and Bennett's (2005) early conceptualization of process tracing and define it as "the analysis of evidence on processes, sequences, and conjunctures of events within a case for the purposes of either developing or testing hypotheses about causal mechanisms that might causally explain the case" (2015, p. 7). The emphasis on chains of events is shared by Collier (2011), who also emphasizes the descriptive component of this approach – noting that the careful description of events and situations at one point in time ("good snapshots of specific moments", p. 824) constitutes the cornerstone of process tracing.

Since the turn of the century, process tracing appears to have experienced a rise in popularity within qualitative scholarship (Beach and Pedersen, 2013). The study of IOs is no exception to this trend – as observed by Checkel (2015), literature on international institutions increasingly puts a premium on fine-grained descriptions and mechanism-based explanations. Importantly, and as discussed by Bennet and Checkel (2015), the tenets of process tracing have also been embraced by constructivist scholars – even if the compatibility between constructivism and process tracing had initially generated some doubts given the emphasis of the latter on *objectivity*. According to these authors, whilst reconciling post-modern or interpretive understandings of constructivism with the philosophical roots of process tracing poses a number of (potentially insurmountable) challenges, this is not so much the case when it comes to more conventional constructivist views – i.e., those positing that "there are standards for assessing some interpretations of social life to be superior to others" (Bennet and Checkel, 2015, p. 15).

With doubts on the compatibility between constructivism and process tracing being progressively dispelled, and in the light of its emphasis on *reconstruction* and on the *identification of the causal forces* shaping a given outcome, process tracing appears to be an approach particularly well-suited to the objective of this research – which aims to understand precisely how the SDG4 targets and indicators came into being.

More specifically, this dissertation turns to what Beach and Pedersen (2013) characterize as the *case-centric variant* of process tracing (2013). This modality can be better understood by opposition to the theory-centric variant that appears to be the object of much of the methodological work on process tracing. Thus, theory-centric variants are oriented at building or testing causal mechanisms with the aim of generalizing those to a wider population. Conversely, case-centric variants aim to "explain a particularly puzzling historical outcome" (Beach and Pedersen, 2013, p. 11). Case-centric studies frequently rely on an inductive logic, "work[*ing*] backward to build a plausible theory based on the available evidence gathered from a multitude of different sources" (p. 92), and are oriented towards reaching minimal sufficiency – that is, to account for the most relevant features of the object of study. In this sense, case-centric process tracing exhibits

some common traits with historical scholarship, in that both are oriented towards offering a sufficient explanation rather than to build generalizable theories. Importantly, case-centric and theory-centric approaches differ in their understanding of causal mechanisms. In the case of theory-centric approaches, causal mechanisms are understood as "relatively simple, parsimonious pathways whereby X contributes to producing Y" (p. 12). Conversely, when it comes to case-centric approaches, mechanisms lack a law-like quality and are considered specific to the context. They operate thus as "heuristic instruments whose function is to help build the best possible explanation of a particular outcome" (p. 19). In this respect, case-centric studies constitute a research procedure kindred to what Gerring (2006) has termed *single-outcome studies*. Unlike conventional case studies, single-outcome case studies are those aiming "to investigate a bounded unit in an attempt to elucidate a single outcome occurring within that unit" (p. 707). They are thus highly idiographic in nature – in the sense that they focus on the uniqueness and singularity of the case rather than aspiring to identify regularities or universal causal properties.

### 4.2.2. Making sense of the data: coding interviews

The principles of process tracing discussed above, guided the process of data analysis and, particularly, the codification of the interview transcripts. Thus, I relied on a qualitative data analysis software (Atlas.ti) in order to iteratively develop the index of codes. The process was guided by Deterding and Waters' (2018) *flexible coding approach*. Flexible coding is devised as an alternative to a grounded-theory approach, which Deterding and Waters judged ill-suited to the specificities of contemporary research – particularly given its emphasis on an inductive logic that, in practice, few scholars really observe[73]. Flexible coding relies on a three-stage process, which starts with a first read of the transcripts in order to identify the "main stories" of the data, along with the development and application of a first list of *index codes*, which may correspond to self-evident categories such as the different sections of the interview guide or the different phases in the lifecycle. The ultimate objective of this first coding stage is to gain some familiarity with the corpus of interviews, and to organize data in order to facilitate text location and retrieval. The second stage in the process is the development of a list of *analytic codes,* corresponding to the themes identified by the original research questions, along with some of the emerging insights gained from the familiarization process, or relevant theoretical constructs identified through the review of the literature. The third stage corresponds to the evaluation and refinement of the codes – by examining their theoretical validity and reliability (for instance, by retrieving all the excerpts corresponding to a given theme and making sure that the code has been applied in a consistent manner).

---

[73] Other limitations identified by the authors include the difficulty in applying the principles of grounded theory to large samples of interviews, or in the context of collaborative projects involving a number of coders.

In the context of my research, *index codes* corresponded to the *date* (month, year) and *forum* (strand of the debate, name of consultation or deliberation mechanism) to which a given excerpt referred. This list of codes was applied to the totality of the interviews, regardless of the research objective they contributed to. The decision to use data/forum codes responded to the process-tracing approach orienting the research. Thus, the use of such codes allowed me to classify interview data according to the *sequence of events* that constitute the focus of process tracing. In the case of *analytic codes*, I developed two separate lists, corresponding to the two research objectives and processes to be reconstructed. In relation to the first research objective, analytical codes referred mainly to the motivations and priorities of the participating organizations, as well as to the instances of inter-organizational interaction. In the case of the second research objective, the codes corresponded to the interviewees' preferences relative to harmonization and reporting strategies, their perceptions of the evolving role of the UIS, and instances of inter-organizational cooperation and competition resulting from quantification requirements. These codes made it possible to reconstruct the chain of events through which the SDG4 targets and associated metrics came into being, and the organizational dynamics and structural forces that shaped such processes.

### 4.2.3. Organizing and presenting evidence: a thick description approach

For the purposes of this research, thick description was judged as an appropriate means to organize and present the findings. As a narrative, contextualized and detail-oriented approach, it suits the sequence-focused and mechanism-oriented reasoning style that constitutes the hallmark of process tracing.

The notion of thick description was coined by Gilbert Ryle (1949) and popularized by Clifford Geertz (1973). Rooted in ethnographic research, it was originally used to refer to interpretive work explicitly oriented towards capturing the insider views of the actors involved in a given process or phenomenon. However, and as noted by Merriam and Tisdell (2016), with the passage of time "it has come to be used to refer to a highly descriptive, detailed presentation of the setting and in particular, the findings of a study" (p. 257). This seems to be the product of a certain confusion (or conflation) with the notion of thick analysis as advanced by Coppedge (1999). The notion refers to those works relying on a detailed knowledge of cases and characterized by rich contextualization. Thus, Collier, Brady and Seawright (2004) observe that Geertz's thick description is only one tool or variety of thick analysis – its specificity lying in its emphasis on subjective meaning.

In the following two chapters, aimed at discussing the results of the dissertation, I rely on a thick analysis approach. In order to report the research findings, I make extensive references to the fieldwork (in the form of verbatim quotes) and privilege a narrative, chronological style as a means to capture the dynamic and evolving nature of the analyzed

phenomena, as well as the subjective meaning that my interviewees attribute to such developments.

**Wrapping up**

This chapter has discussed the data-gathering and analytical strategies on which this dissertation rests. It has thus presented the data-collection and data-analysis instruments (namely, the interview guides and indices of codes) and the theoretical considerations informing them. The chapter has also reflected on the specificities associated with the practice of elite interviewing, and the challenges entailed by the online medium on which the bulk of the interview has relied. Overall, the chapter has established and justified the research design upon which the findings discussed over the next two chapters are based.

**Uncharted territory**
**Negotiating priorities and procedures in a loosely-coupled**
**environment**

**Introduction**

The advent of Education 2030-SDG4 represented simultaneously a form of continuity and a departure from previous instances of education goal-setting, most notably, the Education for All (EFA) agenda and the Millennium Development Goals (MDGs). While the negotiation process built on a well-established tradition that had come to be recognized as a procedural hallmark of the UN system, it also entailed a certain discontinuity – both in terms of content and procedure, the new education agenda represented a move away from its prior iterations. Thus, SDG4 was seized upon by various organizations as a means of renegotiating a new set of shared priorities and principles. At the same time, the coincidence of the MDGs and the EFA deadlines, and the discontent generated by the duality of agendas that had characterized the last decade, entailed the fusion of two distinct, decision-making structures. This proved a challenging endeavor, for in practice, it required finding common ground between different (if not necessarily conflicting) expectations and normative understandings with regard to the legitimate *locus* of power. The negotiation of SDG4 thus represented a major effort in consensus-building that was frequently riddled with conflict.

This chapter aims to make sense of the process through which the global education community charted a new, unified agenda, and attempts to gain an insight into how this process was shaped and how it contributed to the transformation of the identity, routines and relative position of the partaking organizations. The process is examined from two distinct entry points. The first entry point is the process of renegotiation of the very architecture of the debate, i.e., the process through which the MDG and the EFA deliberation mechanisms were merged into a new unified structure. The second entry point is the negotiation of the priorities of the education-for-development realm. To explore this question, the chapter focuses on the intersection of the post-2015 debate and a broader process of normative change, namely, the consolidation of the so-called learning turn.

The chapter is organized as follows. The first section offers a descriptive account of the deliberation and negotiation architecture, set up by the UN in order to craft a new development agenda. This is followed by an analysis of the process through which a consensus was reached on the very structure of the education goal-setting process, i.e., the process through which the venues, working modalities and negotiation practices were renegotiated. The final section of the chapter addresses the process through which the SDG4 debate was instrumentalized by certain actors as an opportunity to position the

improvement of learning outcomes as a priority in the development field, and explains how such a process was resisted by strategically-located constituencies – and how these circumstances forced a certain tweaking of the new norm, in order to secure consensus.

## 5.1. Setting the scene: a bird's eye view of the post-2015 process

The formulation of the Sustainable Development Goals (SDGs) was the result of a long and complex process characterized by its open-ended, participatory quality, as well as by its poly-centric nature. Thus, unlike prior goal-setting exercises, the SDGs did not emerge from a single negotiation chamber, but were the product of a variety of parallel processes that fed into one another. While, as a result of the EFA legacy, the negotiation of the new education targets relied on an architecture of its own, the broader negotiation structure, facilitated by the UN, is central in explaining the process through which SDG4 came into being. Therefore, for contextualization purposes, this section aims to provide an overview of the different initiatives that contributed, to a certain degree, to the shaping of the 2030 Agenda for Sustainable Development.

The section offers a detailed description of the two main strands of the debate, namely, the post-2015 debate and the Rio+2020 process. For each of these, an overview of the multiple initiatives that fed into the workstream is provided, and their interlinkages discussed. The section ends with a discussion of the last stage of the process in which the two strands converge, and offers a description of the follow-up mechanisms, established to monitor the implementation of the new agenda.

### 5.1.1. The legacy of the MDGs: all hands on deck to define the post-2015 agenda

The origins of the SDGs need to be traced back to the so-called *post-2015 process*, a multi-sited debate born out of the need to follow up and give momentum to the MDGs, while simultaneously discussing the development framework. This framework was intended to supersede the MDGs once the 2015 deadline was reached. The process was formally initiated in the context of the 2010 High-Level Plenary Meeting of the UN General Assembly, devised as an MDG summit. In the outcome document, resulting from the meeting, the UN General Assembly not only reaffirmed its commitment to the realization of the MDG agenda, but also requested that the UN-SG take the necessary steps to ensure the continuity of the UN's development agenda in the post-2015 era (UNGA, 2010). Therefore, the outcome document closed by noting:

> We request the Secretary-General to report annually on progress in the implementation of the Millennium Development Goals until 2015 and to make recommendations in his annual reports, as appropriate, for further steps to advance the United Nations development agenda beyond 2015. (UNGA, 2010, p. 29)

As discussed by Granmo and Fourie (2021), the decision to provide a certain degree of continuity to the UN's development agenda (by finding an appropriate substitute for the MDGs) should be understood in connection with the mixed (and sometimes contradictory) sentiments and attitudes surrounding the MDGs that prevailed within the UN quarters by the early 2010s. By 2010 it was increasingly apparent that many of the aspirations advanced by the MDGs could not possibly be realized. Two key reports, published in the run-up to the 2010 summit, were emphatic that the slow and uneven progress that had characterized the first decade of the MDGs rendered it unlikely that the goals would be fulfilled by 2015 – unless an effort was made to accelerate improvement (see UN Secretary-General, 2010; and UNGA, 2010). Thus, while the goals were not deemed unattainable, the limited success of many of them led to an overall sense of pessimism. At the same time, this lack of success did not lead to a questioning of the potential of global agendas and summits – but to the conclusion that, precisely because the MDGs were likely to remain an unfulfilled promise by the 2015 deadline, it was necessary to devise a *new* development agenda. This agenda should be able to continue to inspire the development community to sustain or redouble its efforts. In coherence with a dynamic, well-established within the UN quarters (cf. Chapter 2), the limited success of global agendas and goal-setting exercises somehow reinforced the appetite for such instruments.

The request made to the UN-SG to start working on the post-MDG era gave way to the establishment of several initiatives that coalesced into came to be known as the post-2015 debate (or post-2015 process). Importantly, at the time that such initiatives kicked off, the endpoint of the post-2015 process was far from clear. With a myriad of initiatives unfolding simultaneously, the rapport and connection between the different workstreams remained unclear – even to those partaking in such efforts. The decentralized and loosely-structured nature of the debate arrangement was apparent in the documentation published during that period and has in fact been explicitly highlighted in the main insider accounts of the process published to date – see for instance Kamau, Chasek and O'Connor (2018) and Dodds, Donoghue and Roesch (2017). In this sense, the making of the SDGs represents a clear departure from the *modus operandi* that had given way to the MDGs. Whilst, as discussed in Chapter 2, the Millennium Declaration and the MDGs were crafted by a limited circle of experts and UN bureaucrats, the post-2015 process was devised as a much more participatory and organic process. The process was thus open to a greater variety of inputs, lacked a clear center of gravity and resulted from the simultaneous unfolding of multiple initiatives.

The first initiative, set in motion as a response to the UN General Assembly's request in 2010, was an inter-agency mechanism – the UN System Task Team on the post-2015 UN Development Agenda (UNTT). Established in September 2011[74], and bringing together

---

[74] Note, however, that the UNTT was formally launched in 2012.

senior experts from a wide array of UN agencies and other IOs[75], the UNTT was tasked with the provision of "system-wide support to the post-2015 consultation process, including analytical input, expertise and outreach" (UNTT, 2012a, p. 12). The UNTT was chaired by UN/DESA and UNDP, and while its role was to act as a consultative platform, it soon proved instrumental in defining the overall narrative of the process, and in identifying the overarching principles that were to inform the post-2015 agenda. The first output, issued by the UNTT, *Realizing the Future We Want for All - Report to the Secretary-General* (UNTT, 2012b), had a decisive impact in at least two areas. Firstly, on procedural grounds, the report insisted on the need for a more inclusive process – the underlying narrative being that the technocratic nature of the MDGs explained some of their pitfalls and the limited political traction they had enjoyed in practice. With reference to the MDGs, the report thus noted that: "A more inclusive process might have led to a better understanding of the need to adapt the global goals and targets to country contexts, thus avoiding the unintended perception of the MDGs as a set of uniform targets to be pursued by all countries, regardless of their initial conditions" (UNTT, 2012b, p. 8). On this basis, it called explicitly for "an open and inclusive consultation process in order for the post-2015 agenda to have the best development impact" (UNTT, 2012b, p. 39). Such demands for inclusivity were instrumental in that they legitimized the multi-pronged strategy adopted by the UN-SG when setting in motion a variety of simultaneous initiatives. Furthermore, these demands contributed to ensuring that the open, participatory structure would be maintained over the years to come.

Secondly, and also crucially, the framing provided by the UNTT also contributed to shaping the *ethos* of the new agenda in the making. The calls for an *expanded* agenda (as opposed to the more limited framework, advanced by the MDGs) and for a holistic approach, revolving around the human rights/equality/sustainability triplet, are central in explaining the broad and universalistic nature that would eventually characterize the SDGs. Importantly, the work of the UNTT also laid the foundations of a post-2015 vision that, in the future (see section 5.1.2 below), would enable the reconciliation of two different global agendas. These were the development agenda (embodied by the MDGs) and the sustainability agenda (as advanced by the outcome documents of the Earth Summits of 1992 and 2002 – namely, Agenda 21 and the Rio Principles, and the Johannesburg Plan of Implementation). Thus, in discussing the principles needed to orient the post-2015 agenda, the UNTT identified four highly interdependent core dimensions – namely, environmental sustainability, inclusive economic development, inclusive social development, and peace and security (2012b, p. 24-25). By placing such concepts at the center of the post-2015 debates, the work of the UNTT was central in the ascendance of

---

[75] More than 50 organizations were represented in the UNTT, including the totality of UN agencies and programs, the World Bank, the IMF and the WTO. Organizations were represented by focal points appointed by the principals or leaders of each institution (UNTT, 2012c). An exhaustive list of the participating entities can be found in Annex 1 of the UNTT (2012b).

*sustainability*[76] and *inclusivity*[77] as two meta-norms that would substitute poverty-eradication and human development – the meta-norms that had brought the MDGs together (Granmo and Fourie, 2021).

The contributions of the UNTT were not confined to the 2012 report, as the group continued to provide further input throughout the post-2015 process – for instance through the 2013 report *A Renewed Partnership for Development* (UNTT, 2013). Similar to the report *Realizing the Future We Want for All*, this second publication was also oriented towards shaping the tone of the global conversation, rather than influencing the exact content of the agenda in relation to specific areas. However, the work of the UNTT also proved crucial in relation to particular thematic areas, including education, through the organization of a series of thematic consultations, and through the production of up to 29 Issue Briefs[78] (published between April 2013 and March 2014) aimed at informing the deliberations of another workstream – the Open Working Group (see section 5.1.2 below).

As highlighted by Dodds, Donoghue and Roesch (2017), it is important to bear in mind that the Issue Briefs and the thematic consultations were instrumental in shaping the final set of targets. Thus, the stock-taking exercises and the support labor performed by the UNTT proved to be useful opportunities for UN agencies to shape the new agenda. Such mechanisms enabled IOs to participate in (or at least) affect most of the workstreams that would eventually feed into the SDGs. In other words, the UNTT (and, subsequently, the Technical Support Team, TST), offered UN agencies a privileged outlet to influence the post-2015 process in an indirect way. Simultaneously, and as noted by Kamau, Chasek and O'Connor (2018), the UNTT and the formulation of the Issue Briefs also proved a useful strategy in keeping the particularistic interests of agencies at bay. The UNTT

---

[76] To be sure, the notions of *sustainability* and *sustainable development* had been informing the development conversation since at least the start of the 2000s, and both of these concepts enjoyed considerable normative strength (MDG7 was directly concerned with environmental sustainability, and the Millennium Declaration referred explicitly to Agenda 21). However, it was in the context of the post-2015 debate and the Rio+20 debate that such notions acquired a central character, to the point that they became the "overarching objective" of the development agenda (see Granmo and Fourie, 2021, for a detailed discussion of the parallels between the poverty-eradication and the sustainable-development meta-norms).

[77] Encapsulated by the adage "Leaving no one behind", *inclusivity* would also become a ubiquitous term within the post-2015 debates. As in the case of sustainability, the notion was already widely accepted within the UN quarters by the early 2010s, and the unequal distribution of the MDGs' success had become a talking point in the publications assessing the MDG experience. However, it was not until the start of the post-2015 debate that inequality issues became a core preoccupation within development circles (Granmo and Fourie, 2021; Dodds et al., 2017).

[78] See UNGA (2014a) for a compendium of the Issue Briefs. It should be noted that the briefs were produced by the so-called inter-agency Technical Support Team (commonly known by its acronym, TST) established *under the umbrella* of the UNTT, but formally distinct from it. The TST was thus equally co-chaired by UN/DESA and UNDP, but brought together a more limited number of organizations – around 40 according to the last tally provided by the UNGA (2014a).

provided an institutionalized venue to contribute to the debate, thus discouraging lobbying tactics that would have greatly hindered consensus-building.

A second mechanism put in place by the UN-SG in response to the 2010 High-Level Plenary Meeting of the UN General Assembly was the High-Level Panel of Eminent Persons (HLP), tasked with a rather ambiguous mandate – namely, the provision of recommendations and advice on the development framework beyond 2015. Constituted in July 2012 and made up of 27[79] personalities (mostly, world leaders and political figures, but also civil society representatives and experts), the panel was co-chaired by President Susilo Bambang Yudhoyono of Indonesia, President Ellen Johnson-Sirleaf of Liberia and Prime Minister David Cameron of the United Kingdom. While the composition of the panel was not free from criticism (cf. Dodds et al., 2017), existing accounts note that the HLP eventually became an authoritative, credible source of advice. The final report, submitted by the HLP to the UN-SG, *A New Global Partnership: Eradicate Poverty and Transform Economies Through Sustainable Development* (UN, 2013), proposed a number of transformational shifts[80] that contributed to securing the *aspirational* character of the post-2015 agenda. The report aimed to guarantee that the level of ambition, exhibited by the new development framework was considerably higher than that demonstrated by the MDGs. While accounts on the inner workings of the HLP are notoriously sparse, Dodds, Donoghue and Roesch (2017) recall that the expansive view spearheaded by the HLP owes much to the contribution of representatives from developing countries.

As reported in different studies (see Dodds et al., 2017; or Granmo and Fourie, 2021, *i.a.*), the HLP report did in fact became one of the most influential inputs that shaped the target proposals subsequently advanced by other *fora*. Thus, the HLP coined a number of influential expressions and mottos, most notably, the formula "Leave no one behind", which would eventually become a shortcut for the principles of equity and inclusivity (Thérien and Pouliot, 2020). More importantly, it proposed a set of 12 goals and targets that substantially expanded the scope of the MDGs, while simultaneously reaffirming a fundamental belief in the potential for concise lists of global objectives – albeit with a tweak. Thus, while the report was explicit in establishing that goals should be universal (i.e., applicable to the universe of countries), it also noted that targets should be adapted to national realities (see Appendix 1 in UN, 2013).

---

[79] According to the original Terms of Reference, the Panel assembled 26 eminent persons and a 27th member – the Special Advisor of the Secretary-General for Post-2015 (Amina J. Mohammed), acting as an ex-officio member of the HLP and serving as a liaison to the UN system (UN, 2012). The distinction was, however, effaced with the passage of time, with most of the accounts referring to a 27-person panel.

[80] More specifically, (*1*) "Leave no one behind"; (*2*) put sustainable development at the core; (*3*) transform economies for jobs and inclusive growth; (*4*) build peace and effective, open and accountable institutions for all and (*5*) forge a new global partnership (UN, 2013).

2012 saw also the establishment of a number of consultative mechanisms and public hearings oriented towards collecting the input of national stakeholders, the civil society and sector-specific constituencies. These mechanisms included up to 11 thematic consultations[81], 88 national consultations and a global on-line survey (UNDG, 2013a, 2013b). The three strategies were part of a broader project, launched by the UNDP[82] (the so-called *Building the Post-2015 Development Agenda: Open and Inclusive Consultation* project), in response to the UN-SG's calls for the post-2015 process to adopt "an inclusive, open and transparent process with multi- stakeholder participation" (Zuijderduijn, Egberts and Krämer, 2016).

Thematic consultations were organized around a list of topics, determined by the UN Development Group (UNDG). While the synthesis report warned against interpreting the list as a proxy for those areas deserving their own standalone goal (UNDG, 2013c, p. 54), the selection was eventually perceived as a decision of political significance (or at least an accurate prediction of the status and saliency of certain themes within the UN quarters). Each of the thematic consultations combined on-line engagement modalities on the basis of a common platform[83], various preparatory meetings (regional dialogues or consultations with specific stakeholders), and at least one high-level meeting, oriented towards reaching a consensus – that is, at crafting a common vision around the means and ends of development in each of these areas.

National consultations were led and facilitated by specialized UN agencies, and benefitted from the sponsorship or support of different countries[84] (UNDG, 2013c). In the case of national consultations, these were facilitated by UN country teams, generally with the engagement of the national government, but also in coordination with other stakeholders (private sectors, civil society, etc.)[85]. As noted in the synthesis report, such consultations were driven by the explicit goal to reach "those who would not normally have the possibility of contributing to the post-2015 debate" (UNDG, 2013b, p. 44). To this end, guidance for the organization of national consultations encouraged the use of a

---

[81] Importantly, one of these thematic consultations was devoted to education and proved to be one of the most influential mechanisms (or workstreams) in the definition of new education goals. The specifics of this thematic consultation are thus discussed in a detailed fashion in the next section of this chapter.

[82] Operating under the umbrella of the UNDG.

[83] The platform was known as the *worldwewant2015.org space*. The similarity between the name of the platform and that used for the global survey (MY World) was the source of some confusion among interviewees – a dynamic ultimately indicative of the complexity of the arrangement.

[84] The expectation was that, for each area, at least one developed and one developing country would collaborate in such efforts (Dodds et al., 2017). The rule was, however, not always observed – the consultation on Growth and Employment was exclusively sponsored by Japan (UNDG, 2013c).

[85] The selection of countries was aimed at ensuring that different world regions were represented. However, since consultations were to be facilitated by UN teams with a development cooperation program, high-income countries were scarcely represented in the final sample (UNDG, 2013b).

diversity of channels. However, the format and timeline of the national consultations differed greatly across different countries, and so did their success in terms of inclusiveness (UNDG, 2013c).

Thematic and national consultations were supplemented by the so-called MY World survey – a global, UN-led survey[86] that relied on off-line votes (via paper ballots), mobile-phone technologies and on-line forms (UN, 2014). The survey asked participants to rank six areas in terms of priority (i.e., to select six out of 16 "priorities for a better life"). By 2014, the survey had received over seven million votes from 194 countries. Remarkably, for the purposes of this chapter, the category entitled "a good education" was consistently ranked as a top priority across all age groups, education levels, gender and HDI country groups, with the sole exception of individuals over the age of 61 educated to primary level (UN, 2014).

As noted by Fox and Stoett (2016), the prominence of such participatory approaches in the making of the post-2015 agenda needs to be understood in relation to the progressive ascendance of *citizen participation* as an emerging norm in global governance – and, more generally, in relation to the *deliberative turn* which guided democratic theory from the early 2000s[87]. The broad nature of the consultations, along with the effort made by the UNDG to guarantee a minimum degree of inclusiveness and to reach out to marginalized groups, were certainly touted within UN circles as a symbol of the participatory, democratic *ethos* driving the post-2015 efforts. Consultations were thus repeatedly hailed in both external and internal accounts of the post-2015 process, portrayed as proof of the UN's willingness to break with the MDG spirit (donor-driven, expert-led, insensitive to local priorities) (see for instance Kamau et al., 2018; but also the UN-SG's report *A Life of Dignity for All*, presented to the UN General Assembly in 2013 – UN Secretary-General, 2013).

However, and although the aggregate of the different consultation mechanisms gave way to an unprecedented level of public outreach, it remains unclear to what extent such efforts succeeded in infusing the post-2015 process with the democratic, participatory *ethos* pursued by the UNDP through the project *Building the Post-2015 Development Agenda*. While literature on such questions is very much at a nascent stage, available scholarship cautions against an excessively optimistic reading of the UN's outreach efforts in relation to the post-2015 process. This is the case of Sénit, Biermann and Kalfagianni's (2017) essay on the consultative processes set in motion during the

---

[86] Note however that the survey was developed in partnership with the UN Millennium Campaign, the Overseas Development Institute and the World Wide Web Foundation – a collaboration that, according to Dodds, Donoghue and Roesch (2017), was relevant in securing the broad outreach of the project.

[87] Dryzek, Fox and Stoett (2016) refer to the *democratic turn*, in order to capture the process through which "democratic legitimacy came to be seen in terms of the ability or opportunity to participate in effective deliberation on the part of those subject to collective decisions" (p. 558).

negotiation of SDGs[88]. The authors note that such mechanisms have proved insufficient in compensating for the democratic deficit that has long characterized intergovernmental spaces. This was largely the consequence of limited inclusiveness (i.e., the limited breadth of the sampling of participants, its lack of correspondence with the world population) but also of limited procedural transparency, and the fact that the ultimate influence of such hearings continues to be something of a question mark (given the wide array of inputs that informed the post-2015 process). A similar conclusion has been reached by Fox and Stoett (2016), who argue that, while the new consultation channels are a significant step in the pursuit of democratic legitimacy, they are still a far cry from the deliberative ideal to which the UN aspire – for most of the consultative outlets were unable to ensure a sufficient degree of openness and representativeness. Likewise, different studies have identified a demographic imbalance both in relation to the MY World survey and national consultations (e.g., Gellers, 2016; Siddiqui, Friedmand and Nader, 2014) and concur in highlighting the limits of *mega-surveys*, as a strategy for global civic engagement. Quite significantly, similar concerns were raised in the final evaluation of the *Building the Post-2015 Development Agenda: Open and Inclusive Consultation* project. The final assessment noted that representativeness could be regarded as rather moderate (with more than 70% of the votes coming from five countries), although it judged the project a success, given the practical and political limitations of securing a greater outreach (Zuijderduijn et al., 2016).

Likewise, it remains unclear to what extent the views emerging from such processes had an impact on the definition of the 2030 Agenda. To be sure, the impact and influence of such consultations in the definition of the post-2015 agenda appears to vary considerably across thematic areas and consultation modalities. However, an assessment of the three-pronged consultative strategy, conducted by an independent consultancy group, argues that, in general terms, consultations essentially shaped the debate in an *indirect* way. Thus, while there is evidence that consultations played a relevant role in securing the comprehensiveness and concertedness of the 2030 Agenda[89], the evaluation notes that "there is no direct link between the global consultations and the shaping of the 2030 Agenda by the Member States" (Zuijderduijn et al., 2016, p. 21). The influence of consultations, therefore, appears to have been channeled through other workstreams – most notably, the work of the HLP, the TST Issue Briefs and the side-events that paralleled the OWG sessions and Intergovernmental Negotiations[90] in New York (Zuijderduijn et al., 2016). While this *indirect* use does not render the consultative

---

[88] Note that, in addition to the MY World survey, the authors discuss other consultation mechanisms, namely, the so-called Rio dialogues and the OWG meetings, discussed below.

[89] That is, in guaranteeing a breadth and a level of ambition significantly higher than those exhibited by the MDGs, and in facilitating that issue-specific discussions were held with reference to the debates in each area (Zuijderduijn et al., 2016).

[90] See section 5.1.2 below for a discussion of the OWG sessions and Inter-Governmental Negotiations.

strategy useless, it certainly allowed for a considerable degree of selectivity and discretion in the adoption of its key recommendations.

An additional workstream, set in motion by the UN-SG, was the Sustainable Development Solutions Network (SDSN), launched in August 2012 with the objective, "to mobilize global scientific and technological knowledge on the challenges of sustainable development, including the design and implementation of the post-2015 global sustainable development agenda" (SDSN, 2013, p. 13). The platform has been directed since its inception by Jeffrey Sachs (who, as noted in Chapter 2, had already proved a controversial figure as special advisor to the UN-SG on the MDGs) and led by a core group of prominent figures from the business, academic and public sectors (known as the Leadership Council). Much like the HLP, the elite-like, selective nature of the group render the SDSN something of a black-box in relation to the inner workings of the group and its connection to and embeddedness in the broader post-2015 process architecture (e.g., the way in which its work was expected to feed into other workstreams has never been elaborated on). The input provided by this group in the lead-up to the adoption of the SDGs was synthesized in two relatively succinct reports, *An Action Agenda for Sustainable Development – Report for the UN SG* (SDSN, 2013) and *Indicators and Monitoring Framework for Sustainable Development Goals Launching a Data Revolution for the SDG* (SDSN, 2015)[91]. However, available accounts (see, in particular, the overview provided by Kamau et al., 2018) suggest that such reports came to play an important role. The first of these, which contained the SDSN's own version of the SDGs, was widely circulated and frequently cited during the OWG negotiations, discussed below. Its significance lies in the fact that the report was one of the few inputs, calling explicitly for a concise, succinct list of goals. The report clearly stated that, "The SDSN believes that ten is the maximum practical number. Beyond ten, the goals would lose the benefit of public understanding and motivation. We did not find a way to reduce the SDGs to fewer than ten" (SDSN, 2013, p. 36). It should also be noted that the SDSN Director enjoyed considerable ascendance in relation to the UN leaders steering the process. He acted as an adviser on the post-2015 agenda and had private meetings with the OWG Co-Chairs. Given the high-profile of most of the members of the Leadership Council, the impact of such a group cannot be underestimated, although explicit references to its influence remain scant.

---

[91] It should however be noted that the activity of the SDSN has never been confined to informing the definition of the post-2015 agenda – the mission of the network also revolves around the very implementation of the SDGs (along with the Paris Climate Agreement). The SDSN has thus outlived the post-2015 process and exhibits an expansive work program, which includes educational, research and policy activities. See: https://www.unsdsn.org/about-us

*5.1.2. Rio+20: a clear mandate in an uncertain context*

In parallel to the post-2015 process set up by the UN-SG upon request of the UN General Assembly, a second process began to unfold in 2012 – namely, the development of an agenda for sustainable development. The origins of this workstream can be found in the 2012 United Nations Conference on Sustainable Development, held in Rio de Janeiro and also known as Rio+20 or Earth Summit 2012. The meeting was devised as a continuation of the United Nations Conference on Environment and Development (held in 1992) and the World Summit on Sustainable Development (held in 2002). The outcome document, adopted as an UN General Assembly resolution on July 2012 (Resolution 66/248, commonly known by its title – *The Future We Want*), tasked the UN General Assembly with the establishment of an Open Working Group (OWG), led by Member States and with the mandate to formulate a set of SDGs. The OWG was, in turn, designed as the first step of a broader intergovernmental process that would finalize with a period of Intergovernmental Negotiations (IGN).

The establishment of the OWG represented thus the start of what would operate for some time as a parallel track – that is, one concurrent but not explicitly aligned to the post-2015 debate described in the prior section. The OWG was formally launched in September 2012, in accordance with the tight timeframe established by *The Future We Want* resolution, with the appointment of Maria Luiza Ribeiro Viotti (Ambassador for Brazil) as facilitator for the constitution of the group. However, the process soon stalled, for while Resolution 66/248 had clearly established that the OWG was to be integrated by 30 representatives, the UN Member States could not agree on the nominations (the number of volunteers being higher than the number of seats). This led to the establishment of an alternative format that had no precedent within the UN General Assembly – the creation of *shared seats.* This way, the number of Member States represented in the OWG expanded in practice to 70[92], with countries forming a series of *troikas* (or duos) and being encouraged to speak collectively. As noted by both Dodds, Donoghue and Roesch (2017), and by Kamau, Chasek and O'Connor (2018), the transcendence of such procedural changes can hardly be overstated, for they contributed to a greater diversity of input in the process and reinforced the image of inclusivity and diversity to which the whole post-2015 process aspired.

In February 2013, the two Co-Chairs of the OWG were elected – Csaba Körösi and Macharia Kamau, the Permanent Representatives of Hungary and Kenya, respectively. Both would indeed play a key role in setting the "rules of the game" for the months to follow. Importantly, the production of global goals *led by Member States* had no precedent in the UN context. One of the most relevant decisions in this regard was the establishment of a first *stock-taking period* – a year-long process oriented towards building trust and a "common understanding" of the main topics. According to Kamau, Chasek and O'Connor, this warm-up exercise was deemed key to easing the "anxiety and

---

[92] Note that OWG meetings remained open to those Member States not included in the OWG.

mistrust" that had often proved to be an obstacle in the UN context (with divisions emerging along geopolitical lines), and was intended to depoliticize and to "level the playing field" (Kamau et al., 2018; see specifically Chapters 3 and 4). Thus, the OWG process commenced, with eight sessions organized around thematic clusters. The process ran from March 2013 to February 2014 and exhibited a number of signature features, with no precedent in the UN context. One of these was the effort to bring outsider perspectives into the process, by inviting a number of panelists external to UN agencies, and the deliberation of presentations followed by a relatively spontaneous exchange of views. Also importantly, the sessions were informed by the Issues Briefs, prepared by the TST – a move that signaled the first connection between the Rio+2020 process and the post-2015 debate, described in the prior section (UNGA, 2013a, 2013b, 2014a). Additionally, the stocktaking sessions were paralleled by an increasingly rich program of complementary events, including a number of high-level events and more than 100 side-events (sponsored by a variety of constituencies and held during the hours when the OWG was not reunited) (Dodds et al., 2017; Kamau et al., 2018).

Another arrangement introduced by the OWG was the establishment of a procedure oriented towards engaging with civil society and, more generally, with constituencies other than Member States and UN agencies. This was a means of remaining faithful to the inclusive and open *ethos*, emphasized in the first stages of the debate. The arrangement relied on the Major Groups system, established in the context of the 1992 Earth Summit[93], and extended to include a heterogeneous "Other stakeholders" category (comprising volunteer groups, persons with disabilities, philanthropic organizations, etc.). Civil society representatives were thus afforded a certain number of speaking slots in the OWG sessions[94], and were invited to participate in the so-called Morning Hearings[95], a one-hour long meeting between Major Groups and the OWG Co-Chairs, held on the days that the OWG met. The meetings were explicitly oriented towards building trust between the Co-Chairs and the civil society representatives, as well as

---

[93] The Major Groups system was established in the context of the UN Conference on Environment and Development, and formalized in Agenda 21. They were devised as a means of channeling the participation of a wide range of constituencies in development initiatives. Nine categories were established, namely, Business and Industry, Children and Youth, Farmers, Indigenous Peoples, Local Authorities, NGOs, Scientific and Technological Community, Women, and Workers and Trade Unions.

[94] Note however that access to the OWG negotiations was restricted to accredited organizations, such as those in consultative status with ECOSOC. Consultative status is granted by ECOSOC upon recommendation of the Committee of NGOs. A distinction is made between *general* consultative status (reserved for large international NGOs whose area of work cover most of the issues on the agenda of ECOSOC) and *special* consultative status (reserved for NGOs concerned with a more limited number of themes). An additional non-consultative category is the *Roster* status (reserved for those organizations which do not fall into the former categories, or those already enjoying formal status with a specific UN agency) (UN/DESA, 2020).

[95] Unlike the OWG negotiations, the Morning Hearings were open to any organization or individual willing to self-nominate to participate in the meeting. The selection of the participants remained, however, the prerogative of a specifically-designated steering committee (UN, 2020).

offering the latter an opportunity to shape or impact the new agenda. Although the real degree of influence afforded by such mechanisms remains an open question, emerging evidence suggests that the impact of the civil society was limited in terms of issue-framing or in position-shifting, even if they allowed such groups to prevent certain themes from being dropped from the debate (Sénit, 2020; Sénit et al., 2017; Spijkers and Honniball, 2015).

As noted above, the OWG's stock-taking period signaled a first junction point between the post-2015 process and the Rio+2020 process. However, the exact fit of the two processes remained uncertain to many, leading to a sense of indeterminacy that, ultimately, put many Member States in a difficult position (for they were unclear of the implications of making a contribution to overlapping debates whose compatibility was far from obvious), and fostered a climate of suspicion (Kamau et al., 2018). In order to address such concerns, a resolution was approved in the aftermath of the General Assembly's Special Event towards achieving the MDGs, hosted by the UN General Assembly President (UNGA, 2013c). The resolution is considered to have signaled an organic fusion of the two processes, noting:

> Recognizing the intrinsic interlinkage between poverty eradication and the promotion of sustainable development, we underline the need for a coherent approach that integrates in a balanced manner the three dimensions of sustainable development. This coherent approach involves working towards a single framework and set of goals, universal in nature and applicable to all countries, while taking account of differing national circumstances and respecting national policies and priorities. It should also promote peace and security, democratic governance, the rule of law, gender equality and human rights for all. (UNGA, 2013c, p. 4)

On the basis of this stock-taking period, and once the duality of agendas had been solved, the OWG Co-Chairs (along with the Secretariat) compiled the so-called "Focus Areas Document" (OWG, 2014a). This was presented in the context of the 9th session of the OWG (March 2014) and was devised as a first basis for the negotiations, even though there was some insistence on the part of the Co-Chairs that the document *did not* constitute a first draft, but rather a source of inspiration. Indeed, the document identified up to 19 areas of interest, a figure that would eventually be reduced to 17 in the following months. The negotiations also benefitted from a number of additional inputs, including a compendium of existing goals and targets (OWG, 2014b; OWG, 2014c). In the context of the 11th session of the OWG (May 2014), a Working Document, with a tentative list of goals and targets was presented, while in the 12th session (June 2014), the Zero Draft was formally distributed (OWG 2014d, 2014e). Such documents gave rise to the presentation of a number of amendments and statements on the part of Member States and Major Groups, as well as to informal discussions. On the basis of such contributions, a consolidated document (OWG, 2014f) was presented in the context of the 13th session (July 2014). The proposal, containing 17 SDGs and 169 targets, was adopted by consensus.

During this stage, strategies to reach a compromise were manifold, and varied greatly across different thematic areas (cf. Kamau et al., 2018, for an overview). For the purposes of this chapter, it is, however, relevant to pay particular attention to one of the innovations put forward by the OWG, namely, the devolution of drafting responsibilities to the Co-Chairs, along with the deliberate avoidance of "on-screen negotiations" (i.e., the procedure according to which the position of each party is displayed on a screen). As noted in Kamau's insider account, such strategies were oriented towards avoiding the image of a zero-sum game but also preventing the emergence of "wordsmithing" dynamics (with the debate centered on the wording of the text rather than on the substance). It was thus noted that the use of on-screen techniques renders negotiations particularly difficult in that "delegations do not want to see their names disappear" as it "reduces a chair's crucial room for maneuvering, in the privacy of his or her office, to 'tweak' submitted proposals in the interest of producing a more consensual draft text" (2018, p. 145-147).

The OWG phase was followed by a period of IGN, in which the final proposal of the OWG was discussed. David Donoghue (Permanent Representative of Ireland) and Macharia Kamau (Permanent Representative of Kenya) were appointed as Co-Facilitators of the process (UN, 2015a), which started in January 2015 after a period of informal consultations. The environment of the process has been described as "surprisingly constructive" (Dodds et al., 2017, p. 71). The debates focused on a wide range of issues that went beyond the wording of the goals and targets, including the final declaration, and also the embeddedness of the SDGs with the broader development architecture. One of the key questions in this regard was the formal separation of the SDG debate and the Financing for Development (FfD) negotiations.

It is important to bear in mind that one of the main outcomes of Rio+2020 had been the establishment of an Intergovernmental Committee of Experts on Sustainable Development Finance (ICESDF), expected to provide input to the Third International Conference on Financing for Development (to be held in Addis Ababa, Ethiopia, in July 2015). However, the connection between the SDGs debate and the ICESDF deliberations had long been a question mark, with the committee negotiations relying less than the OWG on a participatory dynamic. The IGN process was thus used to clarify the relationship between the parallel tracks of the SDG and the FfD. This proved, in fact, to be one of the most contentious issues for, in practice, the implementation of a number of the SDG proposals (in particular, the targets relative to the means of implementation and Goal 17) depended largely on the volume of resources afforded by the new FfD agenda, then in the making. Since the final outcome of the FfD process (the Addis Ababa Action Agenda, known as AAAA – UN, 2015b) was perceived by many as disappointing, there was limited enthusiasm in relation to the possibility of integrating references to the AAAA in the declaration text accompanying the SDGs. Eventually, a decision was made to recognize the centrality of the AAAA in the declaration, without attaching the text (Dodds et al., 2017; Kamau et al., 2018).

*5.1.3. A final push: the establishment of follow-up mechanisms*

Intergovernmental negotiations came to an end in July 2015 after three sessions, dedicated exclusively to the finalization of the outcome document. The consolidated declaration, baptized *Transforming our World: The 2030 Agenda for Sustainable Development*, was adopted by the UN General Assembly on 25th September 2015, in the context of the United Nations Summit for the adoption of the post-2015 development agenda (UNGA, 2015). In addition to the SDGs and the principles, commitments and common vision underpinning them, the adopted document also provided detail on a number of review and follow-up mechanisms, that were to play a central role in the years to come. At the global level, review and oversight tasks were placed under the responsibility of the High-Level Political Forum (HLPF), the creation of which was in fact mandated by the Rio+2020 outcome document. The HLPF would operate under the auspices of the UN General Assembly and UN DESA, and their task was to be supported by a newly produced annual progress report on the SDGs prepared by the Secretary-General in cooperation with the UN system.

While the declaration *Transforming our World* discussed the review and implementation strategy in detail, it should be born in mind that by 2015, the monitoring infrastructure was not quite finished. Indeed, the centerpiece of the follow-up strategy, namely the global indicators framework, would remain under discussion for another 18 months. Such a disconnect between the goals and the indicators might come as a surprise in light of the importance attached to the measurement efforts in the context of the post-2015 debates. Indeed, as early as 2014, the UN-SG called for a "data revolution" in a press release, prior to the appointment of an Independent Expert Advisory Group on the Data Revolution for Sustainable Development (UN News, 2014). The group produced a widely-circulated report, *A World that Counts*, which noted that "Data are the lifeblood of decision-making and the raw material for accountability. Without high-quality data providing the right information on the right things at the right time; designing, monitoring and evaluating effective policies becomes almost impossible" (Data Revolution Group, 2014, p. 2). The report heralded the measurement efforts as the lynchpin of the post-2015 agenda, arguing that data improvement was imperative if the new goals were to be realized.

The delay in the release of the global indicator framework needs in fact to be understood in connection with the division of labor, established in 2014 in the context of the OWG and the IGN negotiations – a separation between the political and technical work that, once again, represented a significant departure from the blueprint established by the MDGs. Thus, in December 2014, the identification of the global indicators became the responsibility of the UN Statistical Commission (UNSC)[96]. The UNSC consists of 24

---

[96] It should be however noted that as early as 2013, the UNSC established a number of mechanisms in anticipation of the statistical demands of the post-2015 agenda then in the making. Initially, such efforts were oriented at supporting the IGN process. To this end, a Friends of the Chair group, supported by the UNSD, was established in March 2013. While this group essentially had a monitoring and advisory role, it played a key role in laying the foundations of

Member States, elected by the UN Economic and Social Council for a four-year term, and defines itself as "the highest decision-making body for international statistical activities" (UNSC, 2021). It essentially operates as an intergovernmental forum, the annual meetings of which are attended by a large number of country statisticians (many more than those representing the 24 Member States in office) as well as statisticians of other UN agencies. The organization is served by a permanent secretariat, the UN Statistical Division (UNSD). It is important to bear in mind that the role of the UNSC is one of technical support, guidance and coordination. Thus, while the UNSC is responsible for overseeing the statistical activity of the UN, assisting in the development of new data-collection and harmonization technologies, and advising countries and specialized agencies, the UNSC is not expected to develop indicators of its own (Merry, 2017; Ward, 2004).

In coherence with this mandate, the UNSC adopted a supervision and leadership role in the development of the global framework of indicators for the monitoring of SDGs. Thus, while the UNSC compiled a preliminary list of indicators on the basis of the suggestions and recommendations of a number of UN agencies, it soon delegated the development of the indicator list to an ad-hoc forum, established in March 2015 – the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs). The IAEG-SDGs was made up of representatives of the national statistical offices of 27 Member States (with seats distributed evenly among the UN's world regions). Members were elected for a two-year term[97]. IOs and regional organizations enjoyed an observer status; CSOs, NGOs and other experts were allowed to attend physical meetings and were invited to provide input when deemed appropriate (ECOSOC, 2017).

The negotiation arrangement set up by the IAEG for the development of the global framework gave rise to an intricate sequence of internal and external consultations and plenary meetings. The process reached a milestone in March 2017 with the endorsement of the global indicator framework by the UNSC in its 48th session, and ended formally with the UN General Assembly's adoption of the framework in July 2017 (UNGA, 2017). However, efforts in the development of indicators and data-production did not stop with the adoption of the framework. Not only does the work of the IAEG continue to date, but its work is, in fact, complemented by the contributions of two additional bodies, established by the UNSC – the High-level Group for Partnership, Coordination and Capacity-Building for statistics for the 2030 Agenda for Sustainable Development (HLG-PCCB), and the World Forum on Sustainable Development Data. The former is composed of a number of representatives from national statistics offices, whose role is to provide advice and strategic leadership in the implementation of the so-called data

---

the SDG measurement approach. Thus, the preparatory work conducted by this group (in the form of a compendium of statistical notes transmitted to the Co-Chairs of the OWG, as well as a survey of available data at a country level) ended up shaping the statistical efforts that followed, in a decisive way.

[97] Thus, the membership of the IAEG was updated in June 2019.

revolution, thus supporting the capacity-building, coordination and funding efforts, necessary for strengthening global and national statistical capacity (ECOSOC, 2017). The latter, in turn, was devised as a periodic encounter, oriented towards fostering innovation and cooperation in the production of global data (Data Revolution Group, 2014). Such developments are indicative of the growing centrality acquired by quantitative indicators in the SDG context – data-collection and statistical capacity-building indeed remain among the few areas in which global coordination efforts are as lively today as they were in 2015.

## 5.2. Not the destination but the journey: competing decision-making structures in the lead-up to SDG4

As suggested by the overview provided in the previous section, the process of negotiation of the SDGs was characterized by its open-ended and improvisatory nature. Far from being a master-planned process guided by an explicit and neat roadmap, it retained an open quality well into the very end of the negotiations in 2015. A multitude of processes fed into the making of new goals, from global consultations to advisory panels, including IGN and expert contributions. The negotiation of SDGs was thus multi-layered and anchored to a wide variety of *fora*, venues and arenas – to the extent that, for many organizations, the construction of SDGs became an exercise in multitasking.

The open, polycentric and improvisatory nature of the process has been touted as proof of the horizontal and inclusive nature of the negotiation of the SDGs. This is particularly the case of insider accounts – see for instance Dodds, Donoghue and Roesch (2017) or Kamau, Chasek and O'Connor (2018). Similarly, more theoretically-driven commentators have also highlighted the distinctly inclusive and participatory nature of the process, especially in contrast with its precursors, namely, the MDGs (cf. Biermann and Kanie, 2017; or Thérien and Pouliot, 2020 [98]). Thérien and Pouliot (2020) explicitly draw attention to the bricolage or patchwork quality of the SDG process, and the fact that the negotiation of the new set of goals was informed by a wide variety of practices (from high-level meetings to worldwide surveys) that turned the SDG debate into "a model of inclusive global public policymaking" (p. 11).

While insightful, such accounts tend to neglect (or at least downplay) the fact that the coexistence of multiple, and sometimes overlapping negotiation processes, was not always harmonious or conflict-free, but was affected by power asymmetries and competing expectations as to which are the truly legitimate and determinant negotiation venues. This is the case as the different goal-setting practices that fed into the SDGs were,

---

[98] However, see Browne (2017) for a less celebratory account of the inclusive nature of the process. According to the author, such complexity is one of the key factors that may be blamed for the duplicative character, thus turning SDGs into a self-defeating exercise.

in fact, attached to different constituencies and *loci* of power, and relied on different repertoires of action and varied approaches to consensus-making.

In the education realm such dynamics were particularly apparent given the fact that, since the turn of the century, the global education agenda had been informed by two separate sets of goals – namely, the EFA goals, established in Dakar (WEF, 2000) and the MDGs. Importantly, each of these sets of goals were associated with a specific decision-making architecture and with different communities of practice. Therefore, as noted in Chapter 2, these two agendas emerged from the interaction of different groups of actors, who relied on particular consensus-making scripts. While the EFA agenda (and especially, the so-called Dakar goals) was very much the product of a consensus, carefully crafted by the global education community (i.e., specialized circles of experts and transnational bureaucrats with an education background), and reflecting the multiple priorities of education agencies while allowing civil society to make a meaningful contribution, this was not obviously the case with regard to the MDGs.

Since the negotiation of the SDGs was approached by different agents as an opportunity to put an end to the duality of education agendas, crafting a new set of education targets meant, in practice, combining the decision-making procedures specific to the EFA architecture with the myriad of negotiation processes, set in motion within UN circles as a continuation of the MDGs. This entailed the need to find common ground between different (if not necessarily conflicting) expectations and normative understandings regarding the legitimate *loci* of power, and regarding the acceptable ways of negotiating consensus – that is, the profile and number of those invited to key forums, the nature and formalization of interventions, the sources of knowledge deemed sufficiently authoritative to inform the debate and the appropriate strategies to reach compromises. Thus, some of the major and most persistent disagreements (or misunderstandings) that emerged during the negotiation of SDG4 did not revolve around the very content of the goals, but on the decision-making structures and deliberation scripts on which their negotiation relied.

This section identifies and analyzes some of the key issues around which such tensions crystallized. While the negotiation of the wording of the SDGs has been discussed elsewhere (see especially Moriarty, 2019; and Wulff, 2020), there is a more limited understanding of the process through which a consensus was reached on the very structure of the negotiation process, i.e., the venues, working modalities and negotiation practices deemed legitimate for global- and target-setting purposes. Thus, rather than scrutinizing the way in which each one of the SDG4 targets came into being, this section addresses the process of construction of the very decision-making architecture on which the negotiation of SDG4 was founded.

*5.2.1. Putting EFA to bed? An ambiguous mandate for UNESCO*

As previously mentioned, it is important to consider that, by the turn of the century, there were two sets of global goals addressing education themes, both with a deadline of 2015 – the MDGs (specifically, goals 2 and 3) and the EFA goals, adopted in the context of the Dakar meeting in 2000. The alignment between the two agendas was, however, limited. As recalled by Unterhalter (2007), the two sets of goals differed widely in their level of ambition – the maximalist approach of the EFA contrasted greatly with the minimalist approach of the MDG agenda. More specifically, while the EFA framework articulated a broad vision of education progress, the MDGs made a case for basic schooling and gender parity[99].

The overlap between the education-focused MDG and the EFA goals was eventually settled in favor of the former. Access to education remained the priority, primarily in relation to primary education (Ito, 2012; King, 2007), although access to secondary education eventually gained prominence, as high levels of gross primary enrolment were secured (Lewin, 2008). Thus, in practice, the MDG discourse progressively took precedence over the EFA agenda in a variety of development circles. The EFA rhetoric was thus "largely unfulfilled in practice" (Ahmed, 2014). Also, governments in many low- and middle-income countries focused their resources on the improvement of enrolment levels (Nicolai, Wild, Wales, Hine and Engel, 2014). The unequal ascendance, visibility and political traction enjoyed by different education goals and targets is captured in the overview provided by Mundy and Manion:

> It is important to note that the most widely and consistently endorsed global goals over the past decade are the goals of universal access to primary education and the achievement of gender equity in elementary and secondary education (again measured in relation to access). The wider goals adopted at the World Education Forum (Dakar 2000) and at Jomtien have not received the same level of attention and commitment from heads of states and international organizations – in particular, goals related to early childhood education, youth skills and adult education had unclear targets and have remained neglected. Another neglected area of concern, on that is gaining in international attention, is the importance of "learning" as opposed to schooling. (2015, p. 62-63)

---

[99] It is not clear why MDGs were so limited in the area of education – hardly reflecting the EFA consensus that, by the year 2000, had been in place for a decade. Unterhalter's (2014b) accounts feature among the few comments of the process – and in fact, the author notes that this remains something of a blind spot. According to the author, "Schooling became the locus of measurement for MDG 2 and MDG 3. A number of commentators (Hulme and Scott 2010; Black and White 2003) point to the links between the targets and indicators incorporated in the MDG framework and the International Development Goals set out by the OECD's DAC in 1996. The 1996 International Development Goal education targets limit the broad vision of education outlined in EfA, including only the narrow visions of universal primary education and gender parity in primary and secondary education by 2015 as goals (DAC 1996, 9–11). Ultimately, these were also the goals included in the MDGs. As yet there is no historical account of the diplomatic and organizational processes that resulted in the narrowing of the EfA and Dakar agendas to the MDG framework on education" (p. 179).

The unequal success met by the MDG and the EFA agendas did not go unnoticed in the global education community. Within UNESCO circles, such asymmetries were attributed to the limited ownership of the Dakar goals exhibited by the remainder of the education agencies – a situation reminiscent of the first years of the Education for All agenda. Different UNESCO interviewees attributed such limited commitment to the Dakar goals not only to the insular, self-referential ethos of the rest of education agencies, but also to the limited success or ability of UNESCO in maintaining the momentum around the EFA Agenda[100]. Therefore, a common view was that, as a result of such dynamics, EFA was perceived as "UNESCO's agenda":

> I think in the initial stage, there was a sense of commitment to inter-agency cooperation. But I think one of the main things UNESCO did not invest enough was in showing the senior leadership in those agencies of the value to them of being part of a coordinated push. So they had to take it forward in areas where they have them individual benefit […] So I think, maybe it was a structural issue, an issue about the EFA committee. The EFA committee, in addition to having the EFA committee had a subset of agency meetings where, at the highest level, they would meet alongside or once a year to commit against the same common target […] but at a practical level, they were not very important. (UN26)

> The MDGs had one goal in education, but it was quite a very narrow goal. It was a goal about universalizing primary education. Whereas EFA was much broader than primary education […] But my own analysis is that, for understandable reasons, the international development community… education and development community referred to the MDG framework and not to the EFA framework. As of 2000, it became the language of international cooperation. So, if the MDGs referred to primary education, nobody would be referring to the EFA, which was not a development agenda, it was purely an education agenda. So, in a way I think EFA became seen over time as a UNESCO agenda, and not as an international agenda which was adopted by so many countries and civil society partners, and international development partners etc. But it became to be seen as a UNESCO agenda and everybody was talking about the MDGs. (UN32)

As a result of these dynamics and since the start of the post-2015 process, the future of EFA had been something of a question mark. Particularly within donor circles, the duality of agendas had been perceived as unfortunate at best (and frequently deemed counter-productive). Since EFA was perceived as a UNESCO-dominated parcel, the remainder of the education agencies had an incentive to bring to an end an agenda that they no longer perceived as their own. Thus, and according to interviewees from both UNESCO (see Quote 1 below) and other organizations (see Quote 2 below, from a UNICEF senior official), when conversations on the post-2015 era took off, UNESCO was pressured to avoid entertaining any plans relating to the continuation or reinvigoration of EFA:

> There was a little bit of power dynamic between World Bank, UNICEF and UNESCO. The World Bank and UNICEF shared the same view that UNESCO with its EFA agenda had been having a parallel agenda. UNESCO had not been walking the line, had not been

---

[100] Such diagnostic remarks are in fact consistent with the explanations advanced by specialized literature. See Chapter 2 for a more detailed overview of the limited effectiveness of the EFA coordination mechanisms.

in line with everybody else, who were referring to MDG- UNESCO has been referring to its own agenda. So, if you want, this was an unjustified accusation on the part of World Bank and UNICEF towards UNESCO. But basically, from some internal discussions which were held at high level, the World Bank and the UNICEF negotiated with UNESCO and [*Inaudible - Made it clear?*] "If you start lobbying for a separate agenda, you are on your own. You know, we need one agenda in which we have education, not a separate education agenda. If you agree to do that, UNICEF and the World Bank have no problem that UNESCO should lead on the global agenda. (UN20).

There was a very real risk that we did handle with UNESCO at a very early stage. I remember having a quite high-level meeting with them and with UNICEF's senior staff. And the risk was that UNESCO would try to forge a parallel track, an "EFA 2 track". And within UNICEF one of the really, a really big objective was not to have two parallel tracks, the post-MDGs and then the EFA 2. And we did a lot of work, behind the scenes, on trying to work with UNESCO to agree that we should be dually pitching for bringing these things together. There were letters that went backwards and forwards at quite senior levels to try to get a senior-level buy into that shared direction, so that the teams could work together. But of course, UNESCO is a huge organization and there have been lots of Member States, particularly from certain regions, who had less interest in the MDGs and who were interested in the EFA, particularly the ones that are permanent reps [*short for representatives?*] in Paris and go to the meetings. There was a real political dynamic around that. (UN25)

However, as the second excerpt suggests, there were strong pressures within UNESCO to give continuity to EFA beyond the 2015 deadline. As early as 2011, UNESCO's Executive Board formally requested UNESCO Director-General to start a dialogue on the future of EFA (see section II/3 in the document, *Decisions adopted by the Executive Board at its 186th Session* – UNESCO, 2011). The mandate had been reiterated in 2012, when the Director-General submitted the Strategic vision and plan for UNESCO's advocacy efforts for EFA at global, regional and subregional levels to the Executive Board. In reference to the global-level advocacy strategy, it was noted:

UNESCO will use a series of high-level events in 2012 to kick off a renewed advocacy effort for EFA. UNESCO's contribution to the Rio+20 Summit has helped to reaffirm the importance of education to sustainable development. The launch of the United Nations Secretary-General's new global initiative on education, which seeks to raise the profile of education and generate additional funding for the sector, will enhance the momentum (September 2012) will be the Executive Secretary of the high-level steering committee guiding the initiative […] The first GEM will accelerate this renewed advocacy drive by setting the collective agenda for accelerating progress on EFA up to 2015 and by agreeing on a strategy for shaping the post- 2015 agenda on education. (UNESCO, 2012, p. 2-3)

As a consequence, UNESCO had limited room for maneuver in relation to the future of EFA. Reflecting on the multiple expectations placed on the organization, a UNESCO senior official recalled, "We have been told over and over again that there shouldn't be parallel agendas, and there should be only one global agenda, and there should be only one set of goals, etc., which of course as such makes sense, except that we had a mandate to coordinate EFA up to 2015. And we have been given a mandate from our member states as early as 2011 asking UNESCO to start consultation with the member states on a post-2015 EFA agenda" (UN6). In addition, certain interviewees noted that UNESCO's

unwillingness to abandon the EFA agenda altogether was also a response to the lack of clarity on the breadth and ambition that could be expected from the post-MDG agenda, in the making at that time. Certain UNESCO education circles perceived that it was an organizational mission to safeguard the EFA agenda (holistic, aspirational) against a potential new iteration of the MDGs (narrower, deemed under-ambitious).

Thus, during the 2011/2012 period, UNESCO's activity focused on the preservation and continuity of EFA as an agenda deemed clearly aligned with UNESCO's mission. Conversely, UNESCO's engagement with the global, UN-driven conversation remained more limited. Some interviewees contended that during the early stages of the post-2015 process, UNESCO something of a dormant, disengaged agency, and were thus under the impression that UNESCO had made little effort to affect the outcomes of the UN's discussion on the continuity of the MDGs and/or the adoption of a new agreement revolving around the sustainability themes. The reasons for such limited involvement appear to be largely related to the undetermined, haphazard nature of the post-2015 workstream initiated by UN entities, and the limited insight or prescience exhibited by UNESCO in relation to the UN dynamics. This was not only a result of UNESCO's self-referentiality, but was also a reflection of the complexity and indeterminacy that characterized the early stages of the post-2015 debate. As put it by a CSO organizer:

> I think a thing that the Education sector messed up with… UNESCO was really unorganized and they failed completely to get their… to have any influence over what was going on in New York. I think UNESCO… the kind interpretation is that they just misread the situation and assumed that there would be more space for the UN agencies to come in with technical clarifications and so on at the later stage in the process. The not so kind interpretation is that they actually just failed to even try to read the situation that they weren't strategic enough in their involvement to even think about when to enter. (CSNGO1)

Thus, the engagement of UNESCO with the UN-led process did not take shape until 2013, when UNESCO became the co-lead, along with UNICEF, of the thematic consultation, organized under the auspices of the UNDP as part of a broader consultative strategy (see section 5.1.1. above). The UNESCO/UNICEF shared leadership did create some tension, particularly as the coordination of the process was initially placed under the exclusive leadership of UNESCO. The move was perceived as unjustified within UNESCO quarters – not only because, formally speaking, UNESCO remained the lead education agency, but especially since in 2011 UNESCO itself had initiated a consultative process regarding the future of EFA. A report prepared by UNESCO's Internal Oversight Service noted critically that:

> [*As of 2011*] UNESCO started holding regional consultations on the future of the education sector […] and administered a survey to collect Member States' feedback on the possibility of organizing a World Education Forum 2015 […] UNESCO staff made repeated efforts to convene other development partners to discuss the future of the EFA agenda. *Unfortunately, the response from the other partners in 2011 was not as prompt as UNESCO would have expected.* Despite the Member States' willingness to have UNESCO lead the post-2015 consultations (see 186 EX/Decisions), none of the EFA Co-

conveners, for instance, designated a focal point to the EFA task force set-up by UNESCO to facilitate the thematic consultations. *Interestingly, co-conveners contributed more promptly to the broader UN-sponsored Thematic Consultations held as of 2012, initially under the UNICEF leadership and later placed under a joint UNESCO-UNICEF coordination.* (UNESCO Internal Oversight Service, 2016, p. 109, emphasis mine)

Interestingly, UNICEF interviewees did not appear to consider these dynamics as a significant source of tension. According to a UNICEF senior official, the decision for UNICEF to take the lead in the consultation exercises was essentially a response to UNESCO's budgetary restrictions:

> Quite frankly, UNESCO should have led all the consultations because that's their role. But ah, UNICEF had more capacity and more flexible finance, and paid for most of it. No, I mean a lot of the sort of the underlying stuff was our budgets! UNESCO does a good job of convening and they do a good job of putting their name on that, but they don't necessarily have the staff and the money to throw into it. So that was a very practical partnership. Of course, UNICEF said we want dual leadership with you, if we're picking up the bill. We're not subsidiary [*laughter*]. (UN25)

The consultation ran from September 2012 to March 2013. As for the rest of the thematic consultation, the education debate relied on a combination of face-to-face meetings and on-line consultations supported by the World We Want Platform (UNICEF/UNESCO, 2013). However, one of the key particularities of the education thematic consultation is that it was organized on the basis of a pre-existing architecture – the EFA structure. Thus, the consultation relied on four EFA Regional Meetings[101], an EFA Side Meeting (held in Paris in October 2012), a meeting of the Collective Consultation of NGOs on Education for All (also held in Paris in October 2012) and a Global Meeting (held in Dakar in March 2013)[102] (UNICEF/UNESCO, 2013).

According to certain interviewees, the decision was in response to UNESCO's willingness to reassert itself and reaffirm its authority and leadership capacity, vis-à-vis UNICEF. However, others note that this was not necessarily a strategic move, but simply a decision that proved unexpectedly helpful in bolstering UNESCO's credibility:

> Through the thematic consultation, first of all, we said, "Well, we have a process in place". And I don't think that any other SDG had such a process and structure in place, none of them. So first of all, we needed to signal that. And secondly, that was also an initial victory in the sense of strengthening UNESCO's role and enabling us to continue to lead. It was in the discussions with UNICEF and the other agencies when we said, "Well, okay. We're co-facilitating with UNICEF…" but we needed to tell UNICEF, "You know what? Why are we starting something separate and different? We already

---

[101] More specifically, a meeting for the Arab Region (Sharm el-Sheikh, October, 2012), the African Region (Johannesburg, October, 2012), the Latin American/Caribbean Region (Mexico City, January, 2013), and the Asia/Pacific Region (Bangkok, February/March, 2013).

[102] This was complemented by two mechanisms not directly associated with the EFA architecture – a UN Member States briefing (held in March 2013 in New York) and a consultation with the private sector and donor agencies (held in February 2013 in New York).

have the EFA process, and you're a co-convener of the Dakar EFA agenda. So why don't we just continue doing what we do anyway, because we are already planning in 2012 all these regional EFA conferences? And so we just include into these EFA conferences the consultations about the future agenda, in addition to what these conferences usually did, which was monitoring the EFA progress and looking at priorities. Why don't we just do that?" And UNICEF, I mean, what could they say? They said, "Well, yes. Of course" (UN31)

At the same time, some informants have remarked that the "awakening" of UNESCO and the revitalization of the EFA architecture as a deliberation platform owes much to the pressure exerted by a number of CSOs and NGOs, especially, those brought together by GCE and the Collective Consultation of NGOs on Education for All (CCNGO/EFA). The recollections of different interviewees point to the fact that, at least to a certain degree, UNESCO was forced (or externally encouraged) to play a leading role in the post-2015 context. It is also important to bear in mind that, as some interviewees remarked, such pressure was partially exerted out of self-interest, as, for many of these organizations, UNESCO and the EFA structure remained the main, if not the only, entry point to the UN architecture. The revitalization of the EFA architecture allowed for the input of a number of actors that, otherwise, would have had a much more limited voice in the definition of the post-2015 agenda:

> UNESCO was, at one point, at risk of being bullied […] But UNESCO did, in the end, come through, and defended a broader agenda. They obviously had massive self-interest in doing that […] [*but*] it's difficult because I think we often find ourselves needing to defend UNESCO as a space because there's a sort of a member state engagement, there's and a more accountable structure in UNESCO than UNICEF, which of course doesn't have that same organic representative nature and operates much more as a sort of unaccountable […] UNESCO did, I think, in the end, come through reasonably strong. But I do think it was that alliance between civil society campaigners and UNESCO that headed off an attempt to broaden the agenda. (CSNGO3)

> There was a defined EFA architecture. We were well-positioned in the architecture. We were members of the coordination group of the collective consultation of NGOs in EFA, which was recognized as the NGO, what voice, if you like, or mechanism for participation in the EFA processes. And it was participation that spanned the, well, almost 15 years of the EFA framework, no, the period of the [*inaudible*] framework for actions. So we knew the players. We had credibility within UNESCO. We had official status within UNESCO. We were accredited as an NGO with official relations in UNESCO. So we had easier access, again, if you like. But I think apart from that, because we were players from the beginning, I think they listened to us more. They knew we were a real network. So they knew that the voices we carried had weight. And so it was clear that this was a strategic space that we should occupy. (CSNGO12)

Thus, in a rather unexpected way, the UN-driven conversation became an opportunity for the revitalization of the EFA architecture. The move was particularly significant given the fact that, over the last few years, the effectiveness and convening power of the EFA coordination mechanisms had been severely limited – a situation that had led to a complete restructuring of the EFA architecture (cf. Chapter 2). As discussed below, the central role attributed to EFA regional and global meetings in the context of the UN

thematic conference was only the first step in the progressive transformation of the EFA architecture into one of the main negotiation chambers of the post-2015 education debate.

*5.2.2. Securing a standalone goal: Dakar as a rallying call*

The thematic consultation proved a highly consequential episode for a number of reasons that go beyond the revitalization of the EFA architecture. Firstly, and as will be discussed in the third section of this chapter, the consultation was central in the definition of the *content* of the new agenda. Thus, the synthesis report (UNICEF/UNESCO, 2013) contained a proposal for an overarching education goal that would eventually inspire the OWG proposals – for the report directly informed both the TST Issue Brief on Education (cf. UNGA, 2014a) and the final HLP report (cf. UN, 2013). In addition, the report called explicitly for a holistic, ambitious and broad education agenda, driven by a human-rights based approach. Such an expansive vision would indeed become the basis for the multiple proposals formulated in the EFA and the UN deliberations in the years to come.

Secondly, it is important to bear in mind that, within the global education community, the early stages of the post-2015 debate were shaped by a general sense of uncertainty, if not pessimism, regarding education's overall standing on the international agenda. There was mounting concern that education was no longer a top priority within the development realm, and that, as a consequence, it might not be retained as a priority in the post-2015 agenda then in the making. Such preoccupation was largely a response to the sense of optimism that, according to certain commentators, had imbued the international arena and the UN sphere in response to the progress made by the MDG2, perceived as one of the most successful goals (Unterhalter, 2014b). Optimistic claims as to the "success" of MDG2 were, in fact, routinely tempered by education specialists, who noted that enrolment growth was slowing down, out-of-school children were increasingly concentrated in conflict-affected areas (and consequently, harder to reach), and enrolment and completion could not be mechanically equated to learning (see for instance Burnett and Felsman, 2012; but also Unterhalter, 2014b; and UNICEF/UNESCO, 2013). However, such concerns did not always register beyond the education community. Therefore, when the post-2015 process was initiated within the UN quarters (c. 2012), it was far from clear whether education would feature as a standalone priority. This delicate state of affairs was captured in a somewhat fatalistic essay, published by Nicholas Burnett in 2012. In the paper, Burnett issued a warning note, stating that "Education (and human development more generally) has slipped on the global agenda" and concluded "the international education community should pay as much attention to its declining standing among global priorities and try to reverse it, rather than just assuming that there will be a

future for education" (2012, p. 25-26). A respected and well-connected figure within the development community[103], Burnett's words were to be heeded.

Thus, concerns regarding the risk of "losing" education in the next installment of the development agenda had a profound impact on the thematic consultation relating on education – in particular, on the global meeting, held in Dakar in March 2013. As noted by a wide range of interviewees, the meeting was conducted in a climate of general uncertainty regarding the overall status of learning in the global agenda. This was in fact a point emphatically made by some key negotiators representing UNESCO and UNICEF in the early stages of the process, who perceived that such risks had been progressively fading into public oblivion (thus leading to an unfair assessment of the efforts made by leading education agencies):

> So our initial fight with the thematic consultation was only to make sure that there would be a standalone SDG on education. That's where we're coming from! Because now people look back and say, "Well, we don't have this, and the indicators are not good" [*Annoyed imitation*]. But who would have thought, when I go back to this bad situation in early 2012, that there would be even an education goal? Because the initial discussion was about, "We had 8 MDGs and now, let's have 8 or 10 SDGs, no more". And everybody was talking about transversal and this or that. And we thought that there would be some sort of mixed-up social thing that would talk about health, education, and all these other things that nobody wants to pay for. And so the first fight was, "We need a standalone goal. And we need education to be recognized as such while also education is an enabler for all the other goals." And that's what we achieved by proposing something that the High-Level Panel more or less accepted. I mean, they tweaked the wording a little bit, but it was there. (UN31)

> One of the things I used to say when we sort of were in the process of sort of wrapping up around 2015 work was… there was a point where it wasn't clear that there would be space for an education goal because, of course, the initial UN push was "we need to narrow this down", as always. And people were seeing education as a cross-cutting input into other goals rather than a goal in its own right. And of course, every sector had that fear without a doubt […] So therefore we do need to engage in the process because it could run away without us. (UN25)

The risk of "losing" education from the global agenda thus became a rallying call for the global education community – a question around which different organizations could agree despite differences in terms of priorities or target preference[104]. The education

---

[103] A former Assistant Director-General for Education at UNESCO, Burnett had occupied other relevant positions as Director of the UNESCO Education for All Global Monitoring Report and World Bank Human Development Manager for Africa.

[104] Note however that, as discussed in the third section of the chapter, the consultation (and the Dakar meeting in particular) operated simultaneously as a site of normative struggle. More specifically, the consultation prompted (or rendered visible) a polarized debate around the relative importance of learning outcomes (as opposed to a broader approach to education quality). The divisiveness, created by such questions, earned the Dakar meeting the nickname of "Dakar Wars" – a denomination indicative of the fact that, although the meeting was successful in uniting the global education community (and reinforced the appetite to preserve the autonomy of the field),

consultation, and the Dakar meeting in particular, were therefore remembered by different interviewees as a moment of union – an event that reinforced a sense of cohesion and solidarity in response to a (perceived) external threat. Importantly, this perception is shared across different constituencies, as reflected in the views of a CSO representative and a UNICEF negotiator in the following excerpts:

> I think at that time, the education community was still united under the kind of main goal of making sure that there would be a standalone goal on education. I mean all of us in a way had an interest to make sure that Education came across as a really obvious priority with lots and lots and lots of arguments around why it was so central to any new post 2015 agenda. And I think that really helped that group also to kind of summarize the stuff because it felt like the more points we make the better. (CSNGO5)

> At that point there were still some conversations, within countries and within the UN, whether education should be as standalone goal. There was still talk about the social services consolidated goal. And I think that the Dakar consultation, where ministers, academic, the civil society… made it clear that education has to be treated as a standalone goal. And I think that this set the tone on what had to be included. A second reason was that it also said that education is much broader… in a way, it was saying, whatever that comes out, it should represent both the EFA and the MDGs. (UN1)

To be sure, and as the post-2015 process unfolded, the different workstreams set up by the UN-SG issued some reassuring signals regarding the status of learning in the overall agenda. As shown in Table 5.1. below, most of the proposals published during the 2012-2013 period placed education in a prominent position and suggested education-specific goals – even if the proposals differed notoriously in ambition. A similar level of consensus around education themes was indeed highlighted by the Overseas Development Institute, responsible for one of the first efforts of cataloguing and systematizing the proposals for a post-2015 agenda (see Bergh and Couturier, 2013).

**Table 5.1. References to education goals and targets – Key publications 2012-2013**

| Outcome document, year of publication and responsible entity | Education goal | Education targets/Description of the goal |
|---|---|---|
| UNTT<br><br>2012<br><br>*Realizing the Future We Want for All – Report to the Secretary-General* | | Does not include a proposed list of goals but education is education is assumed to continue as a standalone goal: "The objectives captured by the MDGs could be consolidated under the four different dimensions, which would provide continuity, though some goals may need to be defined in a broader sense to capture the global challenges ahead. As an example, the educational goal would go beyond improving access to schooling (quantity) to emphasize improving the relevance and quality of education at all levels" (p. 34). |

---

this process did not entail the erasure of the ideological and normative divisions that had long characterized the education-for-development realm.

| | | |
|---|---|---|
| HLP<br><br>2013<br><br>*A new global partnership: Eradicate poverty and transform economies through sustainable development* | Provide quality education and lifelong learning | - Increase by x% the proportion of children able to access and complete pre-primary education.<br>- Ensure every child, regardless of circumstance, completes primary education able to read, write and count well enough to meet minimum learning standards.<br>- Ensure every child, regardless of circumstance, has access to lower secondary education and increase the proportion of adolescents who achieve recognised and measurable learning outcomes to x%<br>- Increase the number of young and adult women and men with the skills, including technical and vocational, needed for work by x% |
| SDSN<br><br>2013<br><br>*An Action Agenda for Sustainable Development – Report for the UN Secretary-General* | Ensure effective learning for all children and youth for life and livelihood[105] | - All girls and boys have equal access to quality ECD programs.<br>- All girls and boys receive quality primary and secondary education that focuses on learning outcomes and on reducing the dropout rate to zero.<br>- Youth unemployment rate is below [10] percent. |
| UN Secretary-General<br><br>2013<br><br>*A life of dignity for all: Accelerating progress towards the Millennium Development Goals and advancing the United Nations development agenda beyond 2015* | Provide quality education and lifelong learning | Young people should be able to receive high-quality education and learning, from early childhood development to post-primary schooling, including not only formal schooling but also life skills and vocational education and training |

Source: Author's elaboration.

However, the workstream set in motion after Rio+20 did not offer any reassurance that education would be the object of a separate, standalone goal. As discussed in the prior section, the open-ended and slow-moving nature of the process rendered it difficult to anticipate how many goals could be expected. In fact, even if education featured as a separate focus area in the first OWG outcome document summarizing the first year of thematic discussions (and published in March 2014), and subsequently, featured as a

---

[105] Note that, in the amended goal proposal included in the thematic report published in 2014, the education goal and targets read slightly different – "Ensure Effective Learning for All Children and Youth for Life and Livelihood" (SDSN, 2014, p. 95). This chapter refers however to the SDSN original proposal for the revisions did not affect substantially the spirit of the proposal – in terms of ambition or in what concerns the relative status of education within the general development agenda.

standalone goal in the Zero Draft of the OWG proposal (published in June 2014), this did not completely assuage the concerns of the education community. The situation was compounded by the UN-SG's report *The road to dignity by 2030*. In the document, the UN-SG called for a new effort to craft a concise, easy to communicate and action-oriented agenda – a comment followed by the somewhat ambiguous suggestion of rearranging the goals proposed in the Zero Draft, advanced by the OWG:

> Member States have agreed that the agenda laid out by the Open Working Group will be the main basis for the post-2015 intergovernmental process. We now have the opportunity to frame the goals and targets in a way that reflects the ambition of a universal and transformative agenda. I note, in particular, the possibility of maintaining the 17 goals and rearranging them in a focused and concise manner that enables the necessary global awareness and implementation at the country level. (UNGA, 2014b, p. 15)

Against such an uncertain scenario, different interviewees argued that, along with the education thematic consultation, the preservation of education as a standalone goal owes much to the results of the MY World survey animated by the UNDP. As noted previously, education was consistently ranked as a top priority across all age groups, education levels, gender and HDI country groups, with the sole exception of individuals over the age of 61 educated to primary level (UN, 2014). Such an unambiguous success was thus seized by different negotiators as a means of securing a prominent place for education in the development framework then in the making. In the years to come, and particularly once the SDG debate shifted to New York, the results of the survey proved a useful instrument in capturing a certain degree of attention in relation to education issues – frequently seen as of little transcendence among certain diplomatic circles. The (somewhat unexpected) potential of the MY World survey is reflected in the words of a civil society organizer, who noted:

> If you know about the MY World survey, I personally think is the critical thing that got us the education goal. Across all demographics across all countries, education was the first, except the really old people [*laughter*]. Therefore, it kind of became easy for us to ask for our own goal. Like the first sentence one ended up saying was "it's the most important issue for everyone" […] And it subsequently has been a leverage for in New York, when we ask for anything related to education [by saying] "you might remember that we were the most important issue anyone ever asked for so listen to us!" […] As an advocacy took, it was useful. We didn't do anything to influence it, it just came out and we were like, "Wow, this is useful" (CSNGO10)

### 5.2.3. From Dakar to Oman: the conversation bifurcates

As noted in the prior sections, the education thematic consultation operated as a turning point in the definition of the post-2015 education agenda, not only as it contributed directly to secure an education standalone goal to be included in the next generation of development goals, but also as it entailed a revitalization of the EFA architecture. However, while the reinvigoration of the EFA machinery had been widely accepted as a pragmatic strategy to organize the post-2015 education debate, the EFA bureaucracy soon

took a life of its own, thus becoming a source of divisiveness or, at least, confusion. This was a consequence of the fact that, immediately after the WEF 2000, certain key EFA mechanisms and decision-making spaces (and the EFA Steering Committee in particular[106]) continued to work on their own objective for a post-2015 goal – a proposal with a rather ambiguous rapport with that being discussed in the context of the OWG. Thus, with the OWG deliberations in full swing, the autonomy and initiative exhibited by UNESCO and the EFA were seen as problematic or at least, misleading. An OECD official involved in the post-2015 debate summarized the general unease created by the coexistence of two parallel tracks as follows (see Quote 1 below), while a representative from an NGO referred to the existence of parallel tracks as one of the most contentious elements of the debate (see Quote 2 below):

> There was always a kind of… I wouldn't say it was uneasy, but there was a tension between the processes that were going on in New York and the processes that were going on in Paris around the EFA Steering Committee. And it was never very clear, where was SDG4 going to emerge from? Was it going to emerge from the EFA-based Steering Committee, or was it going to emerge from the SDG process in New York? And there was several attempts during the year 2013, 2015 to try and bring the two things together. (LASS3)

> I think that the biggest part of the debate was kind of how the communities would come together, because there were camps that were in favor of the Education for All process being in the lead, and there are other camps that were in favor of the SDG process being the lead process. So I think that was the most controversial issue, there is evidence of a fractured process. (CSNGO16)

The origins of this bifurcation of the post-2015 education conversation can be found in UNESCO's decision to publish a position paper discussing the future of education post-2015[107] (UNESCO, 2014a, 2014b). The significance of this document lies not so much in its content but in its *timing*. The paper came out in February 2014, at a time when the OWG workstream was fully underway. According to UNESCO's interviewees, UNESCO's decision to enter the scene with an agenda of its own was far from straightforward – it resulted indeed from a combination of internal lobbying (on the part

---

[106] See Chapter 2 for a discussion on the origins and mandate of the EFA Steering Committee. For the purposes of this chapter, it might be important to bear in mind the composition of the group, oriented at securing the representativity of a wide range of constituencies (UNESCO, 2016). As of 2015, the list of members included six representatives of Member States (one per world region); three representatives of the E9 Initiative, two representatives of the host country of the next WEF (to be held in 2015 in the Rep. of Korea), five representatives of the EFA Convening Agencies (UNESCO, UNDP, UNFPA, UNICEF and the World Bank), as well as individual representatives of the GPE, the OECD, the private sector (Intel), Education International and two representatives for the civil society (GCE and ASPBAE, with ANCEFA as an alternative) (UNESCO, 2015a).

[107] The move was all the more relevant considering that UNESCO was, in fact, one of the few organizations that had not produced a document of sorts during the early stage of the post-2015 debate – indeed, the production of these types of documents had been a fixture during 2012/2013, with a wide range of organizations publishing their own proposals and comments regarding the post-2015 education agenda (see King and Palmer, 2013).

of certain senior officials, unhappy with the passive role played by UNESCO so far) and external pressures (on the part of the CCNGO/EFA and a number of countries). The following excerpt from a UNESCO senior official represents one of the most comprehensive accounts of the process:

> It has to do with the fact that when we started out in 2012, UNESCO was weak and not seen as a leader. And whenever there's a vacuum, somebody will fill it, right? Especially when it looks like something that could be big and successful. And I think that UNESCO's mistake, not intentional but just because of the circumstances, was that we have not come up on the first of January of 2012 to say, "Okay, guys. We are the leaders, and here's what we're going to do". And so others have started to do things. And the position paper is a good example because there has been hesitation. And we had internal discussions, and there were people in UNESCO who said right from the beginning or at least very early in that year 2012, "We have to publish something. We have to come up with something" […] And then the UN came in through this consultation process on the SDGs and actually condemned all the UN agencies to be facilitators. And when the thematic consultation started, we were told that we were not allowed to take any position whatsoever. And we have courageously carried through the thematic consultations all the way to the end in 2013 and then said, "Okay. So now thematic consultations ended. What happens next? What about the co-facilitators of the consultations?" And the UN secretary said, "Well, you and UNICEF, you have to continue now consultations around the outcomes and help shape this agenda, but you're still the facilitators". And then at some point, we said, "You know what? We are UNESCO. We have a mandate in education. We're the only organization that has a mandate in education. And we have to publish a position paper". And we had also many member states who were pushing who said, "Come on, every NGO… ActionAid has a position paper. Save the Children has a position paper. The Commonwealth has a paper, the Francophone has a paper!" (UN6)

The significance of external pressures was particularly emphasized by certain representatives of the CSOs who remained uncertain UNESCO's genuine interest in playing a leading role in the process. The following description of the process, offered by a CSO representative, is illustrative of this perspective:

> There was this period when basically, UNESCO had practically almost lost… it was 90% surrendered the process as far as defining a clear education agenda within the education constituency to New York. So in September of 2013, they had announced in New York that there would be no EFA anymore, that now the discussions will be around New York, and the whole process is on education would be around New York. And so we lobbied along with… specially using the space of the CCNGO, with governments from France and Japan. These two countries were the ones who were very vocal in an upcoming UNESCO General Conference to ask UNESCO, "What are you doing?!" as far as the new agenda for education is concerned. "The world is passing us by. We have a whole structure. We have processes to debate a new agenda. We have the whole wealth of experience coming out of EFA, and we have EFA. Do you just throw that away?" So they lobbied strongly within the UNESCO General Conference, and we pushed the civil society. And so it was only after that general conference which instructed UNESCO to set in place a set of processes to define a new agenda that it got its act together. And it's through those sets of processes that the Muscat Agreement basically was mooted or was agreed. (CSNGO12)

UNESCO's position paper became the basis of an entire, independent workstream, parallel to that unfolding in New York – for the publication was indeed the first stage in

a three-step sequence oriented at crafting an "EFA input" in the post-2015 process. Thus, the position paper inspired the production of a Joint Proposal of the EFA Steering Committee on Education Post-2015 (UNESCO, 2014c) and this, in turn, formed the basis for the so-called *Muscat Agreement* – i.e., the outcome document of a Global Education for All Meeting (GEM)[108], held in Oman in May 2014[109]. The line of continuity between these three documents is apparent not only in the vision and general principles to which they all refer, but also in the list of proposed goals and targets – which, as shown by Table A.1 in Annex 2, bear a strong resemblance.

The clarity of vision displayed by such documents contrasted greatly with the ambiguity they exhibited in relation to the embeddedness of the EFA line of work within the broader post-2015 architecture. The overall sense of confusion stemmed from the fact that, while these documents were explicitly presented as part of a concerted effort to articulate an EFA-specific vision, they were also incrementally clear on the fact that the duality of education agendas needed to end. Thus, UNESCO's position paper noted that "UNESCO advocates for a single, clearly-defined, global education agenda, which should be an integral part of the broader international development framework" (UNESCO, 2014a, p. 2), and the Muscat Agreement emphasized that "Every effort will be made to ensure coherence between what is agreed in September 2015 at the High-level UN Summit as part of the global development agenda with the post-2015 education agenda approved at the WEF 2015 in the Republic of Korea in May 2015" (UNESCO, 2014e, p 4). Despite their brevity, such observations represent a turning point as they signaled the end of EFA as an independent education agenda[110].

Thus, and somewhat paradoxically, while there was growing consensus (or acceptance) regarding the need to unify the EFA and the UN agendas, the conversation on the *content*

---

[108] As discussed in Chapter 2, GEMs are the mechanisms that, since 2012, have replaced the Working Group and the High-Level Group in the global EFA coordination architecture. On this occasion, the gathering brought together heterogenous groups of some 300 people, including 50 national delegations, representatives from the EFA convener agencies, bilateral and multilateral donors, representatives from the civil society and the private sector, and selected experts (UNESCO, 2014d).

[109] Remarkably, the GEM was originally scheduled to take place in March 2014, before the OWG. Such a timeline would have allowed for the Muscat Agreement to have greater influence on the OWG workstream. However, it has not been possible to discern the reasons behind the change in the GEM dates – it remains unclear whether the delay responds to a strategic choice to avoid intruding or distorting the New York process on the part of the EFA-conveners, although other accounts suggest this might have been the case (see Moriarty, 2019).

[110] To some extent, such calls for a unified agenda can be seen as the formalized expression of a horizon that has long been perceived as inevitable, given the pressures on the part of most EFA-conveners (see section 5.2.1 above). The consolidation of education as a (probable) standalone goal arguably precipitated the widespread acceptance of this scenario. Had education been merged with other social services in the UN-workstream (or addressed as a cross-cutting issue), the continuation of a separate EFA agenda could have been justified. However, once it became clear that education was to become a standalone goal in the development agenda, any prolongation of the EFA goals was more easily portrayed as an ill-advised course of action.

of the new education agenda was bifurcating. Although the three aforementioned papers clearly advocated for a single, harmonized education agenda, they remained unclear as to how this convergence should eventually occur – and especially as to whether the OWG proposal would be inspired by the EFA agenda, or the other way round[111]. The Muscat Agreement closed with a somewhat ambiguous observation in this regard:

> We commit to using this Statement as a reference for the negotiations in the global consultations on the post-2015 development agenda, in order to ensure that this latter has a strong education component. To this end, we ask the Director-General of UNESCO to share this document with all Member States of UNESCO, the Secretary-General of the United Nations (UN), the co-chairs of the Open Working Group, the Committee on Sustainable Development Finance as well as key stakeholders. We further commit to using this Statement for ongoing national, regional and global consultations on the post-2015 education agenda, to be approved at the World Education Forum 2015, which will be hosted by the Republic of Korea in May 2015. Our expectation is that this will be an integral part of the global development agenda to be adopted at the UN Summit in New York City in September 2015. (UNESCO, 2014e, p. 3-4)

The significance of this observation lies in the fact that, by the time the document came out, the OWG had already published a preliminary proposal for the new set of goals and targets, the Working Document for the 11th session of the OWG (OWG, 2014d), followed by the Zero Draft published in June 2014 (OWG, 2014e). Interestingly, such proposals were characterized by their holistic, expansive nature. The breadth of the OWG agenda came in fact as a surprise to some of the EFA Steering Committee parties. According to a Member State representative and an NGO officer:

> Some were afraid that you would lose out on a more comprehensive agenda, that the New York or the UN SDG agenda would be much narrower. But, eventually it turned the other way round […] And this was a kind of a… for me it was a paradox, because I argued at the initial stage that if we would not be very precise, and very narrow, then New York would actually make decisions and make it narrower, but the other way round happened. And that was a surprise. (CR2)

> At one point, you had a progressive agenda coming through the wider UN SDG process and some conservative forces pushing the education side. (CSNGO3)

The expansive nature of the first OWG drafts is attributed to different causes by different interviewees. UNESCO and UNICEF informants were keen to note that the ambitious nature of the agenda was largely the product of the input provided by thematic

---

[111] A similar degree of ambiguity can be observed in the UNESCO programmatic documents, prepared during this period. Thus, one of the nine strategic objectives identified by the 2014-2021 Medium-Term Strategy and approved by the General Conference at its 37th session was "Advancing Education for All and shaping the future international education agenda". The document recalled that "To strengthen UNESCO's participation and comparative advantage in a reforming United Nations system, which is increasingly collaborating and delivering as one, […] UNESCO will continue to provide leadership and/or coordination functions in a number of United Nations system-wide initiatives […] which will shape both the contours and content of the post-2015 development agenda, while remaining fully committed to education for all (EFA) in the field of education" (UNESCO, 2014d, p. 6).

consultations, channeled to the OWG in the form of the TST Issue Brief (see Quote 1 below). Others pointed to the importance of the input from Member States, and to the specificity of the consensus-making mechanism used by the OWG Co-Chairs, explicitly oriented at conflict-avoidance (see Quote 2 below):

> In the OWG countries were supported by what was called the Technical Support Teams, TST. Every sector. So, for education, we had the colleagues in UNICEF and UNESCO. But all the other agencies that have some role in education also attended and could contributed inputs through UNESCO and UNICEF. So in TST calls we would be on the calls, and even the TST had inputs on the targets. Basically UNESCO and UNICEF used what came out of the global and regional consultations, and the Muscat document as basis for providing inputs. So that's why you see some similarities. I think that there are 80% similarities. The differences were largely arising from some of the countries not accepting some of the proposals that were made by the agencies. (UN1)

> I think that, a UN member states led process like this, it is very difficult to take something out that is already included, because then you are in a way in the kind of diplomatic process. You don't, you get a lot of enemies if you try to take out things. So it is easier to add on, So I think that is what happened in New York, and it's not only for education, it's also for most of the other SDG goals, that you are adding on and you come up with a very ambitious and very comprehensive agenda. (CR2)

More in general, it can be argued that the expansionist nature of the OWG proposal was in some ways an inevitable outcome, for the very structure of the OWG deliberations was conducive to *at least* incorporating the goals, targets and commitments reached by other global agendas or conventions. As can be observed in Table A.2 in Annex 2, the inputs that fed into the work of the OWG were often compilations of existing proposals and of the issue areas, addressed by global and regional organizations. Even if the exact level of ambition remained an open question, the breadth of the agenda (i.e., the number of addressed themes) eventually became an inevitability. It was thus a reasonable expectation that the OWG proposals would address a wide range of dimensions and education levels – rather than prioritizing certain issues at the risk of alienating those constituencies that felt "responsible" for certain themes.

*5.2.4. Infusing the OWG with the spirit of Muscat*

Even if the OWG proposal could be judged as satisfactory and sufficiently broad in scope, UNESCO needed to prove that the EFA workstream would not be fully inconsequential or a futile exercise – in order to preserve a sense of organizational credibility. It is also relevant to note that the Muscat Agreement soon become the preferred agenda of a number of NGO and CSO federations – thus, the Muscat Agreement and the EFA-Steering Committee Joint Proposal were explicitly endorsed by UNESCO's Collective Consultation of NGOs on EFA. The Final Declaration resulting from the CCNGO's 7th meeting (held in Santiago de Chile in May 2014) highlighted its explicit support for the Joint Proposal of the EFA Steering Committee and for the Muscat Agreement (UNESCO, 2014b).

Given the need to prove that the EFA deliberations on the post-2015 agenda had not been in vain *without* endangering the eventual convergence of the EFA and the UN agendas, ensuring a close alignment between the proposed education targets of the OWG and those of the Muscat Agreement became an institutional priority for UNESCO. An internal report, prepared for the consideration of the Executive Board in August 2014, observed that UNESCO was committed to ensuring this alignment "through providing comments and analytical inputs to various draft versions of the OWG" (p. 2) and by "inform[*ing*] Member States represented in the OWG on the goal and targets of the Muscat Agreement, via their Permanent Delegations to UNESCO and their Permanent Missions to the United Nations, and through communications, briefings and advocacy events" (UNESCO, 2014b, p. 2). More proactive strategies included the organization of an "Information Meeting on the Muscat Agreement" during the 12th session of the OWG (June 2014). The event was co-hosted by Argentina, Brazil and Norway, with presentations given by the Chair of the EFA Steering Committee (D. Vedeler), the President of the GCE for Education and Vice-Chair of the Committee (C. Croso) and representatives of Education International, UNESCO and UNICEF.

In such efforts, UNESCO benefitted from the crucial support of UNICEF. The collaboration of the two organizations in the thematic consultation appeared to have created the basis for a productive relationship. Certain interviewees noted thus that, despite prior schisms (and a history of limited coordination), the two organizations had collaborated closely[112] since late 2013, with UNICEF acquiring a more prominent and proactive role within the EFA architecture. Such a move was largely made possible by UNESCO's explicit abandonment of any plans to establish a separate agenda and to recreate the EFA/MDG duality in the post-2015 era. UNICEF's support, in fact, proved key in securing the legitimacy of the EFA vision (by preventing it from being perceived as "UNESCO's agenda"). At the same time, UNICEF's support was instrumental in assisting UNESCO in affecting and penetrating UN deliberations – for UNICEF appeared to be much better connected and to have a greater insight into the UN dynamics and the inner workings of the New York negotiating chambers. As put it by a UNESCO negotiator:

> UNICEF, I mean, they have been our closest and strongest ally. And one reason why UNICEF was so important is because I told you of this divide between Paris and New York and because of things going on at the UN in the New York… It might seem strange, but it's really because of the place where things are happening. There's so many things you hear in the corridor or a friend tells you over lunch, "Oh, you know – they're now planning to do this or that". And we were somewhat out of it because our headquarters is not in New York. And so UNICEF was instrumental along the process, especially towards the later part, in giving us information or in being in the kitchen, as we say, when things were decided and in liaising with us for the education agenda. (UN31)

---

[112] Such rapprochement appears to owe much to the personal disposition and brokering capacity of the UNICEF negotiator (J. Naidoo). Remarkably, Naidoo became Director of the Education For All and International Education Coordination at UNESCO in 2015 (a position subsequently transformed into Director of UNESCO's Division for Education 2030 Support and Coordination).

However, the success of the attempts to "infuse" the OWG deliberations with the results of the EFA consensus remains difficult to assess – different interviewees hold different (and sometimes contradictory) views as to the relative influence of the Muscat Agreement on the work of the OWG. Within UNESCO circles, as well as on the part of EFA Steering Committee informants, it is generally assumed that lobbying proved effective and that, during the last period of the OWG negotiations, the education goal and targets in the OWG Outcome Document became much more aligned with those proposed in the Muscat Agreement. According to one of UNESCO's negotiators, the OWG proposal was vetted on multiple occasions by the lead of the EFA Steering Committee and the UNESCO Education Sector (see Quote 1 below). A CSO organizer noted similarly that the wording of the OWG group had been meaningfully altered as a result of the advocacy campaign, orchestrated by the representatives of the EFA Steering Committee (see Quote 2 below).

> Muscat said, "This is what it means to have life-long learning and equal and inclusive quality education for all. And it can't be less". And that had such a strong impact on the Open Working Group. They really aligned, but it was back and forth, sometimes on a day-to-day basis. We were here every evening until 8 or 9 o'clock, a small group in revising and reviewing and drafting our messages to New York say, "Wait, this is what the education community has decided", and arguing in long tables with comments for every single target what it was reflecting and why we couldn't go beyond or underneath, why we could not formulate it in a way that meant less than that. Because one of the principles of the UN process was that we should not have SDGs and targets that were less ambitious than what we had before. And so EFA, of course, was our standard, but we were also looking at other international conventions and other agreements that had been there, and we needed to do all this research and argue like lawyers for every single one of the targets […] It was the outcome of all these consultations that we were holding and the expectation that UNESCO would speak up for that. And that is how the initial SDG agenda and whatever the education goal was at first came closer and closer and looked more and more like the goal that had been identified by the education community. And in the end, they were almost identical. *But it was not us compromising*. (UN6, emphasis mine)

> After the GEM, they came up with a list of targets for the post EFA agenda. But by then the Open Working group had just released its targets. So, therefore, a group of people actually travelled from there to meet the OWG and brief them on what the educational community, including member states had to say... and if you see the trajectory of what the different versions of the language of the targets, there was some change in the targets made as a result of that meeting. (CSNGO10)

Such views were indeed corroborated by one of the EFA Steering Committee Vice-Chairs, who even pinpointed certain specific changes attributable to the influence of the Muscat Agreement. Interestingly, an informant noted that, in order to secure such changes, UNESCO and the EFA Steering Committee had operated largely in an indirect way – namely, by lobbying Member States so that national delegations advocated such modifications in the context of the OWG and the IGN that followed. It appears, therefore, that the EFA Steering Committee relied on a multi-pronged strategy, consisting of high-level negotiations (i.e., direct exchanges between the leads of UNESCO and the OWG) and lobbying and advocacy efforts aimed at securing allies among the Member States:

So you had to try to balance by trying to work within the education community with as much autonomy as possible, to try to get an agenda as good as possible. Knowing that, at the end, you need to link up with what happened in New York. So we had to try to find out what's the possibilities within the education community and then to try to influence the New York process, and to make sure, at the end of the day, to have one [*single agenda*]. And what was crucial from the process I led, where we actually made a difference for the New York process, was in two respects. One was that the word "inclusive", it was inserted in the EFA Steering Committee, and it was followed up by Brazil and supported by others in the New York process. So that was one issue that, it was not might, it was Brazil, and it was a kind of shoulder influence by the EFA Steering Committee to the New York process. The other which I think was even more important is that, in June, I think, the version the SDG version or SDG4 in June did not include a separate target on literacy, and together with my- vice chair [*deleted for anonymity purposes*] we went also with UNESCO to New York and had a meeting with the delegation in New York where we strongly advocated for literacy to be included as a separate target. And it was, in the final version it was included, as you know. And that happened in the July meeting, at a kind of late stage in the process. (CR2)

Importantly, the perception of the process as a collaborative endeavor was not exclusive to EFA circles – some external observers of the process corroborated such views. An OECD observer thus emphasized that the notion according to which the EFA process had been taken over by the OWG was an inaccurate one:

So like I said, there was quite a lot of correspondence between the leaders of the Open Working Group, this guy [*deleted for anonymity purposes*]. He was leading that, and then Mr. Chang, who was leading it from the UNESCO side. So there was a lot of interaction. And like I say, there was one or two targets where the wording was a little bit different to what one group had. But basically, I mean, the short story is that the New York processes superseded the EFA processes, but there was a negotiation between the two, and I think both sides were generally content with what came out in the end. (LASS3)

However, other informants remained more skeptical as to the real influence of the Muscat Agreement on the work of the OWG, and expressed regret or even disappointment at the lack of coordination between the two processes, and the (perceived) lack of foresight on the part of UNESCO:

My sense is that it should have been the 4th or 5th OWG session, when the global education meeting happened. If it had happened like 4 to 5 months earlier, one would have had these outcomes to feed into the beginning of the OWG. Then they were still discussing what should be in the agenda. But then [*by the time the Muscat Agreement came out*] it was too late, the OWG never endorsed to a substantial degree what the GEM result, and in fact it became politically difficult for them to do so, I mean, the OWG was set up to make decisions. So they cannot just way "we are not going to decide for this because stakeholders have decided for us". They produced a couple of drafts, and they did some tweaking to incorporate some key aspects, but things would have been easier if UNESCO had facilitated… if it the schedule was less messed up. I mean, instead of being the 4th or 5th session or whatever it was, it should have been the 1st or the 2nd. Then, we would possibly have had better synergy between the two… but hey, what can we do. (CSNGO4)

In a way, like I actually believe that the rejection of the Open Working Group of the UNESCO attempts to change the language in accordance with the Muscat agreement in June 2014, I don't believe that was a matter of ambition, I believe it was really a matter

of process. The thing with these global level processes is that sooner or later, a process will get a life of its own. And it's suddenly very difficult to change the course of direction without somehow messing up a lot of the other pieces. And I think that's what happened quite early on in the Open Working Group. It's also a social phenomenon in the sense that you have this bunch of people that have been struggling with this language for months and months and month, and they develop a certain ownership, a certain commitment to the process, to the language, to the agreement, to the consensus that they've built. So I think they just felt that they couldn't like, lifting in language from another process would have jeopardized the process and the kind of delicate balance of kind of everybody having compromised on something that you had in the Open Working Group. I mean, with Muscat, you have a process that is fundamentally fucked up, sorry to put it that way. (CSNGO5)

So if you compare Muscat and Incheon… Muscat agreement was more strongly reflecting the view of the education community. But along the way we had to adapt our position to the one of the UN discussion. So at some point our values dropped from the 4.7 statement, the sentence became too long for anybody to remember. So I had to keep my sheet of paper always in front of me! (CR5)

Discerning the degree of influence exerted by the Muscat Agreement on the final list of targets eventually adopted by the UN is indeed a task beyond the scope of this chapter. It should be noted, however, that the first proposal advanced by the OWG in May 2014 certainly underwent a number of modifications in subsequent months, both in the context of the OWG deliberations and during the IGN. As reflected in Table A.3 in the Annex of this dissertation, relevant differences between the multiple versions of the education targets included (a) the inclusion of *relevant* before the reference to learning outcomes (target 4.1.); a more ambitious framing of the target on early childhood development, care and preprimary education (4.2.); an explicit reference to quality and affordable TVET and tertiary education (4.3); an explicit reference to *youth* (in addition to adults) in the reference to literacy and numeracy (4.6); the expansive definition of the notion of knowledge and skills, needed to promote sustainable development (4.7); the disappearance of an explicit reference to *completion rates*; the disappearance of the explicit reference to the *demands of the labor market* when discussing the question of learning outcomes and skills and greater detail on the question of the Means of Implementation, *i.a.*

However, identifying the origins and precise motivation behind each one of these changes is an impossible exercise – not only as most of my informants had a limited insight into the dynamics of the OWG, but particularly because the key role of the OWG in introducing such amendments (and the rejection to do so on-screen) rendered it impossible to identify by who they were introduced. However, Wulff's (2020) detailed account of the OWG deliberations suggest that the changes requested by Member States were typically particularistic in nature (it seems unclear that any country was driven by a well-articulated, comprehensive vision of the education goal). CSO constituencies, however, were more consistent and persistent in their demands, insisting upon the need

for a human-rights based, equity-oriented and ambitious agenda[113]. Overall, determining the influence of the Muscat Agreement remains a virtually impossible endeavor, for, as previously discussed, the promoters of this agreement typically acted *through* Member States or CSO representatives, willing to transfer the EFA vision to the UN negotiation chamber.

*5.2.5. A culture shock? Competing scripts and organizational routines in the negotiation of global goals*

As the section above suggests, the coexistence of the UN and EFA workstreams frequently proved a source of friction and controversy. Interestingly, the tensions caused by the existence of two parallel tracks did not stem from a fundamental incompatibility of the proposals that emerged from each of the strands – at the end of the day, the Muscat Agreement and the first OWG proposals exhibited a notable degree of alignment. Rather, such animosity appears to have its origins in the mismatch between the bureaucratic culture of the UN sphere, and the deliberation routines that characterize the EFA architecture. This, in turn, reflects different understandings of what qualifies as a democratic and inclusive debate. More specifically, a number of CSO/NGO representatives and certain UNESCO negotiators often expressed their discomfort (and sometimes, open irritation) over the negotiation protocols and the deliberation dynamics set in place by the UN process, which they judged inappropriate for the construction of a consensual yet pluralistic education agenda[114].

Such tensions were particularly apparent in the assessment of the open-ended and decentralized nature of the UN-led, post-2015 debates provided by a number of CSO and NGO officers. Representatives from these constituencies routinely expressed a preference for the more formalized and centralized debate structures set up by the EFA architecture

---

[113] As discussed in the next sessions, an exception to this CSO consensus was the status of the learning outcomes agenda.

[114] While the EFA architecture was frequently characterized as a more advantageous structure for the engagement of NGOs and CSOs, it is important to bear in mind that such mechanisms were not free from criticism. The centrality of consultation practices that characterizes the EFA space was deemed insufficient to confer a truly inclusive and participatory nature on the EFA workstream, particularly given the absence of appropriate guidance provided in a timely manner. In this sense, it is important to recall that the decentralized nature of many CSOs and federative NGOs rendered it impossible for the secretariats to make an input in global and regional debates if the various units have not been given the opportunity to discuss the issue under consideration. Such limitations were eloquently captured by a CSO representative in the EFA Steering Committee, who noted "The documentation was released late and there wasn't time to be able to comment. Sometimes we'd be given one week. And you know to consult, for example, a membership of 35 organizations was very difficult […] Sadly, I think the IT infrastructure is also one of the limiting factors, especially in many of the developing countries, and the other thing from [*name of the region*] perspective which I saw as a problem, was that the consultations were, in most of the cases, done in English" (CSNGO15).

and UNESCO's CCNGO/EFA mechanism. As the following excerpts suggest, these views were consistently held among various representatives of the CSO sector:

> I think the problem in some ways is that whilst civil society needs these formal spaces for engagement and we need the speaking slots as part of an interactive dialogue with the member states to influence the agenda and blah blah blah, private sector and the bilaterals don't need that in the same way, and they have a much more direct access to policy-makers. (CSNGO5)

> For example, the EFA processes, and defining the new Framework for Action for Education 2030, these were the processes that were clearly laid out. You knew what was going to happen next. You could fight for who gets to make the decisions, and what was the logic in these bodies making those decisions. So there were arenas to engage, to begin with. There was no such thing very clearly in the UN. Even in the meetings, we had an advantage in terms of participation in meetings, because we also had official NGO status with ECOSOC, and at least gets you through the front gate, no? […] But even if you get into the events, most of the conversations were closed door, no? And unlike other events where there was a clear way by which civil society can participate, certainly in education. (CSNGO12)

As discussed below, much of the criticism of the UN-led process, voiced by CSO and NGO constituencies revolved around the role of the OWG and the IGN negotiations. It is however important to bear in mind that the remainder of the mechanisms set up by the UN-SG were also a frequent object of criticism, for most of them relied on engagement modalities (expert-driven, celebrity-centered, with participation made conditional on personal invitations, etc.) with which such organizations had little experience or they perceived as inappropriate to build consensus. For instance, in relation to the Global Education First Initiative (GEFI)[115], set up by the UN-SG, and the designation of UK Prime Minister Gordon Brown as the UN Special Envoy for Global Education, a CSO organizer observed, "We were also worried about GEFI, we were concerned it was becoming the epicenter, that the leadership was moving to New York. And Ban Ki-moon had invited Gordon Brown, he was the UN Envoy for education… this was very important, eh? Because the Special Envoy deconstructs the role of the Special Rapporteur for Education. So we felt this was a fight for the leadership, for the architecture. And this is why in our meeting in New York we decided to invite Kishore Singh [*Special Rapporteur on the right to education*], and yes, on purpose, we did not invite Gordon Brown" (CSNGO2). Similarly, in relation to the HLP, a CSO organizer described the group as "bit of a free-for-all public forum where everyone has their own say, and not

---

[115] GEFI was launched in 2012 by the UN Secretary-General and revolved around three main priorities, namely put every child in primary school, improve the quality of education, and foster a sense of global citizenship. The initiative was explicitly oriented towards raising the political profile of education and increasing financial support for education, and brought together a number of development partners – including UNESCO, UNICEF, the World Bank, GPE, EI, GCE and the UN Special Envoy for Education. GEFI put forward a number of advocacy and mobilization activities, including a Youth Advocacy Group and the designation of ten "Champion Countries" committed to increase the visibility of the initiative and catalyzed political and financial support for education.

necessarily a lot of stuff coming out; a big jamboree, without necessarily it being gendered into anything concrete on any given thing" (CSNGO10).

Thus, and somewhat paradoxically, the multi-layered structure of the debate was not necessarily experienced as a pathway to inclusion but rather as an obstacle. In fact, and as captured in the quote below, the general sense of indeterminacy and the multi-sited nature of the SDG process was perceived by some informants (and particularly, by those enjoying a more peripherical position, or a more tenuous rapport with the UN galaxy) as a *deliberate* intent to prevent grassroots organizations from engaging in a meaningful way or to make an influential contribution. Regardless of their veracity, such misgivings are informative in their own right – for they denote a limited sense of belonging or ownership of the post-2015 process:

> And one more thing that should be mentioned, the process was transparent. There were plenty, plenty of meetings, plenty of chances for civil society to participate. The consultation process was huge: online, live, WebEx, Skype, and so on, and so on. But let me tell you something […] I think this was on purpose, to make it so complicated. Many, many, many social partners, especially civil society in smaller NGOs, they could not follow. I mean, to understand the process, the consultation process, it was so complicated. You need a master's degrees in minimum to understand what's going on […] you don't know under which conditions you could participate, and where to register, and there were so many deadlines. It was hysteria. It was crazy. We were running after the deadlines and the web pages and meetings and so on. *I believe it was done on purpose.* (CSNGO13, emphasis mine)

In addition to the decentralized nature of the debate, another recurrent source of concern was the background of the Member States' representatives participating in the UN-led process. Such preoccupations featured in fact as one of the most cross-cutting sources of concern – the difficulties caused by the absence of OWG negotiators with an education background were brought up by a number of interviewees[116]. Such dynamics are ultimately indicative of a certain eagerness to preserve the autonomy of the global education field (as a sphere of *specialized practice*, governed by a distinct set of norms), operating as a unifying force around which organizations with different agendas coalesce:

> It was a bit, I think, contradictory, because even in the New York, in the OWG discussions, all the countries were not represented by education people. They were represented by Foreign Officers or development people. So you know, sometimes in

---

[116] According to different interviewees, one of the most egregious consequences of the centrality of outsider (i.e., non-education) perspectives in the context of the OWG negotiations, was the inclusion of a target centered on scholarships (see Target 4.b., UNGA, 2015). The target appears to have emerged from the pressure exerted by a number of the G77 countries, while benefitting from the complicity of high-income countries (that saw the target as a "lower-hanging fruit", given that a substantive part of their ODA budget is already targeted at scholarships). While many informants noted that the proposal was received with skepticism on the part of the education community (who perceived that the target was likely to exacerbate brain-drain dynamics), the concerted efforts of national delegates to preserve it rendered its suppression impossible (UN11, CSNGO8, UN14).

regional meetings we had a different message that from country representatives in New York. You know, the regional meeting, attended by Ministries of Education, they were seeing universal secondary education was not realistic… but at the New York level, most people there were coming from the ministries of international development or foreign affairs… they were in agreement that we need this ambition, that the world is changing. So that was a bit strange. (UN26)

There is a problem with the process here. And the problem is Member States in Paris [*in reference to UNESCO's permanent delegations*], and the opinion of member states in Paris, is not necessarily the opinion of the same Member States in New York [*in reference to the delegates participating in the OWG deliberations*]. So that, I would say, coherence or consistency of the positions were not a reality, and it explains more the problems of member states rather than the problem of UNESCO. (UN3)

In response to such dynamics, and as previously noted, various organizations represented in the EFA Steering Committee, and which perceived the Muscat Agreement as the "real education agenda", set up a number of strategies to influence Member States' representatives participating in the OWG discussions. However, it soon became evident that different agencies were unequally equipped to play a relevant role in the OWG process – as a consequence of different resources, access and overall credibility within the UN environment. At one end of the spectrum, and as noted above, UNICEF advisers featured as the most able negotiators, largely as a consequence of their efforts to cultivate a sense of proximity with national delegations in New York (a move, in turn, made possible by the geographic location of the fund but also on account of pre-existing connections). The ease and adeptness exhibited by UNICEF negotiators in their "indirect" use of Member States is captured in the excerpts below:

Within the UN system there's a certain foothold that the member states have to play. But if you use the big meetings… They were held in every single group and, [*deleted for anonymity purposes*] went to pretty much every single one of them. Then you're building up the constituencies of Member States in a particular direction, who then feeds into the negotiation. It's not about technical agencies using their voices. It's about working with the member states to get some sort of agreement through consultation, which they then take into the negotiations. (UN25)

Do you know what influenced OWG decisions? As I said… UN agencies ourselves could not have direct influence, because we are observers. We could provide technical inputs, but the political decisions were much more important. So what we did was, using the EFA process we had direct connection with Member States… right? So, using EFA member states we tried to influence the OWG decision. So we found some allies which had similar thinking over many things… so, for instance, England, Norway, I think Sweden. I don't remember all the details but also we used different countries to push different agendas […] Our role is basically providing and advising inputs from a purely technical point of view, but we really had to use political means. And the only way we could use political pressure in the decision-making process is using member states. Like in some of the OWG meeting sessions, we literally approached some member states representatives just before the meeting! And we provided like an informal memo and asked them to say… something which we want to push. (UN14)

Sitting in the middle ground, UNESCO lacked UNICEF's familiarity or confidence with the UN environment, but could afford to resort to an indirect outreach strategy – a rather

diffuse advocacy campaign, oriented at disseminating the Muscat Agreement rather than courting the vote of specific Member States. Complementarily, and as documented in the excerpt below, UNESCO reached out to the national delegations in New York using the Paris delegations as a *liaison*, and relied on UNICEF's negotiators as a means of gaining an insight into the OWG mechanics and tempos.

> One challenge that I think we mastered quite well was to bridge this gap between Paris discussions and New York discussions and to make sure that those representatives of Members States who were in meetings and negotiations in New York would be well-briefed about the education process and the outcomes at every single step, which went from, well, the UNESCO position paper through this initial paper of this EFA Steering Committee and then the Muscat Agreement. And so at every step we have worked like mad to publish these documents, to get them translated into all languages, and to get them out to all the networks including civil society so that they could help in at least communicating these to the people in New York to say, "Watch out. There are other people from the same country or government who have actually signed up to this", or talk to them and see… And we have even at one point facilitated UNESCO delegations in Paris to attend meetings in New York of the SDG committee to take forward the positions just to bridge this communication gap there was. (UN31)

At the other end of the spectrum, the capacity to engage in the UN negotiations was considerably limited for certain CSOs and federative NGOs. The main reason for this lies in the OWG accreditation policy, as access to OWG negotiations was, in practice, restricted to organizations in consultative status with the United Nations Economic and Social Council (ECOSOC) – a circumstance that complicated the engagement of some relevant EFA constituencies representing the civil society sector[117]. This was compounded by the absence of a Major Group dedicated to education[118] – for, as discussed in the prior section, Major Group Organizing Partners played a central role in coordinating the interventions of accredited organizations. This meant that, particularly in the case of heterogeneous and large Major Groups (as is the case of the NGO Major Group), participation opportunities for education constituencies remained limited. Thus, largely as a result of the accreditation and participation rules, civil society networks, social movements and grassroots organizations perceived it as virtually impossible to

---

[117] Even if it did not render it impossible. As discussed in the prior section, OWG hearings relied on a more open, participatory invitation policy; any actor (collective, individual) could self-nominate him/herself to participate in the sessions. However, these spaces proven among the less influential of the whole OWG process (Sénit, 2020). In addition, since it corresponded to a Steering Committee to select a number of speakers from the pool of self-nominated participants, well-connected organizations with experience in the New York arena again enjoyed considerable advantage in making their voices heard.

[118] Note that the Education and Academia Stakeholder Group, presently integrated in the UN Major Groups and Other Stakeholders system, was not formalized until 2016. The group has since proved to be an important outlet for the participation of the education civil society community in the follow-up and review of the SDGs. Its organizing partners are the Global Campaign for Education, the International Council for Adult Education, Education International and the European Student's Union, therefore, including some of the organizations that found greater difficulty in engaging with the OWG deliberations.

make a meaningful contribution to the OWG deliberations. The difficulties encountered by these organizations are summarized in the words of a CSO and an NGO organizer, who noted:

> There was no stakeholder group dedicated for education. So you really had to go through existing stakeholder groups to make your case, no? So, that meant building bridges, strong bridges with the stronger stakeholder groups like in this case the women, the trade unions. And it was largely the women and the trade unions that we were building alliances with to get the education agendas through, although we attempted also through the youth. It was no use, see? As this required being able to travel to New York all the time, having a person in New York, which was well beyond our capacity. (CSNGO10)

> [*In reference to the Major Group mechanism*] This was a structure that was a legacy, a legacy from Rio, and on the on hand it had a disadvantage, as all the NGOs are grouped under a larger group, and the NGO group is the broadest, and this left NGOs with little room... And it's not only that we had little room, it's also the pressure of having to find a common position. NGOs represent very different interests and positions, and this is why it was not working. Because there was a very strong of group of environment NGO and they wanted to focus on environment. (CSNGO9)

Access difficulties, and the unease caused among the non-governmental sector by the complexity of the post-2015 architecture, were compounded (or reinforced) by the generalized perception that informal exchanges, interpersonal connections, camaraderie and agreements behind closed doors (or what an informant defined as "backdoor engagement" – DON4) played a major, and frequently determinant role in the negotiations. The centrality of informal spaces was perceived as a major impediment among those organizations with no prior experience in the New York arena – many representatives of which noted critically that they did not have the time or resources to develop the connections necessary to exert a meaningful influence. Such perceptions are clearly articulated in the quotes below from a representative of the philanthropic sector and a CSO organizer, respectively:

> The thing with this kind of negotiations is they are held over a period of time and it's not just one event where it happens. There's a series of events and series of negotiations that happen, that result in each one of the goals and the indicators underneath them, and unless you're willing to invest that amount of time into it, you won't have an impact. So it's a matter of resource allocation. There are organizations, intergovernmental organizations, that do this on a full-time basis as their job. And so we try to build relationships with those kind of people but that's all. (PRI2)

> I mean, I think all of it was decided in the corridors. And that's exactly why it was such a limitation that we didn't have the resources to be present. Whereas if you look at [*name of a large CSO, deleted for anonymity purposes*], for instance, that had [*name of designated negotiator, deleted for anonymity purposes*] essentially fly into all the meetings, and mingle, and network with all the people that made the decisions, that's what makes the difference. (CSNGO8)

At the same time, the coexistence of an EFA and a UN debates even put under strain some well-connected organizations, for it essentially required that advocacy efforts be duplicated and resources (human, economic) devoted to two separate *fora*. In practice,

many organizations decided to focus on one particular strand of the debate, on the basis of their access and capacity to play an influential role. Interestingly, certain organizations that perceived themselves as sharing a number of ideological affinities ended up engaging in a tacit division of labor, as suggested by the quote below:

> Our decision was in terms of capacities, and obviously access, was that with EFA we would be more effective in strengthening an education voice and making sure there was a strong education agenda which was broad-based and which drew consensus widely in the education constituency. And let that be the force that then influences the SDG process. Instead of us, in a way, killing ourselves, in knocking on doors in New York, or in the very vague processes that were emerging then on the SDGs. So that was the first call […] we made a clear choice that our main game was Paris, but we would support organizations who are active and who are better placed to carry the education flag in New York. (CSNGO12)

An exception to such dynamics was the case of a specific organization affiliated to a global union federation – and, consequently, able to operate through the Major Group system. This organization also benefitted from an early engagement with the UN-led process. It featured, in fact, as one of the few education organizations that, since the start of the post-2015 debate, decided to privilege the input of the UN branch, largely as a result of the lobbying efforts of some of its officials. Such early involvement allowed the designated negotiator to cultivate a network of affine delegates and envoys from other organizations, as well as develop a certain ability to code-switch (i.e., to adapt ones' language and message to different targets). Such an outreach strategy eventually proved crucial as a result of the importance of informal negotiations.

> I think we entered the process at an early stage compared to many other civil society actors, and that meant that we could for instance be one of the keynote speakers at that kind of early session of the Open Working Group, when we were still kind of framing or trying to find the scope- define the scope of the issue. I think we benefited in some ways from the fact that there were so few civil society organizations that work explicitly on education […] So we ended up coming across as the organization working on education […] and that gave us access to a lot of actually, meetings, a lot of background information and a whole kind of network of different civil society actors, that I think really helped us in our work. (CSNGO1)

In general terms, however, the UN process was largely perceived as exceedingly burdensome, onerous and challenging among CSO and NGO constituencies. As a consequence, a number of non-government representatives highlighted the risks of a (perceived) shift in the *locus* of power – from UNESCO to the UN or, as stated by certain interviewees, from Paris to New York (CSNGO12, CSNGO10, CSNGO2). The following excerpt by a CSO organizer summarizes these multiple sources of concern, as well as the sense that the *locus* of power has definitely shifted to the UN in what concerns the definition and monitoring of global agendas:

> There are established ways of working that are getting dropped while the new ones don't come up. The shift to New York is not without its risks. Because education isn't exactly-doesn't exactly have history in New York. There is a climate change community both in

civil society and in terms of member states, but nobody knows or cares about education issues in New York. We are used to talking to UNESCO – we know UNESCO and we like UNESCO, but UNESCO doesn't necessarily call the shots in this process […] There is something taken for granted in terms of the EFA Steering Committee, and so and so, and the CCNGO focal point. People will know exactly what you are. Whereas even civil society stakeholders [*in New York*] don't really understand our jargon, they don't understand our language. I have been doing applications for the Open Working Group for our members, and I had to take an application from one of our members and rewrite it, removing all the education terminology and putting in New York terminology. We don't speak the same language. We don't share the same set of priorities. But again, there isn't a lot of choice on the matter. That's how it is and that's how it will be for the next 15 years, and that's why we just need to figure out how to do it. It's kind of "time to wake up and smell the coffee" – UNESCO remains important but we need to also start working with New York. (CSNGO10)

### 5.2.6. Back into routine: crafting a Framework for Action

As the prior section suggests, the format and participation modalities imposed by the UN workstream (and particularly by the OWG and the IGN) risked having a detrimental effect on the sense of ownership of the education community, in relation to the SDGs. This was particularly the case with regard to certain CSOs, NGOs and national decision-makers brought together by the EFA global and regional mechanisms. This appetite for an education agenda *specific* to the global education community is behind the efforts to *complement* the outcomes of the SDG process with two additional artefacts – the World Education Forum 2015 (WEF 2015) and the Framework for Action (FFA). At the same time, UNESCO had a formal commitment with Member States to give some continuity to the EFA agenda. One of the key resolutions of the 37th session of UNESCO's General Conference included an invitation to the Director-General to "facilitate the debate and continue to consult Member States and stakeholders in the development of the global objectives and targets as well as the development of a 'framework for action' for education post-2015, including through the existing global and regional EFA and MDG Coordination mechanisms, and regional consultations" and a request to UNESCO to "carry out efforts to ensure that the global education conference, which will be hosted by the Republic of Korea in the spring of 2015, will result in concrete recommendations and an approved framework of action on the post-2015 education agenda" (UNESCO, 2013a, p. 10).

The WEF 2015 was held in Incheon (Rep. of Korea) in May 2015. Organized by UNESCO in close collaboration with a variety of IOs, with a mandate in education (UNICEF, the World Bank, UNDP, UNFPA, UN Women and UNHCR), it brought together some 1,600 participants, representing a wide variety of constituencies (including 120 Ministers and a number of heads and administrators of international agencies and development partners). The main outcome document of the WEF 2015 was the Incheon Declaration-Education 2030 – a succinct document reaffirming the need for a broad, aspirational and ambitious agenda and, more importantly, entrusting UNESCO with the mandate to lead and coordinate the SDG4-Education 2030 agenda (VVAA, 2015a). It

also signaled a line of continuity with the 1990 WCEFA and the WEF 2000 meetings previously held in Jomtien and Dakar, while confirming the final convergence with the SDG architecture.

The FFA, in turn, was devised as a "roadmap" to guide the implementation of SDG4. The final version of the text[119] was adopted by 184 countries in the context of the 38th session of the General Conference of UNESCO, held in November 2015. The FFA was oriented towards elaborating on the vision of the SDGs in relation to education (i.e., unpacking the principles and rationales underpinning the new education goal) and providing guidance on the implementation strategy, along with the coordination, financing and monitoring mechanisms deemed necessary to realize the new education agenda. Importantly, it confirmed the continuity of the CCNGO and the EFA Steering Committee (the latter, under the form of a renewed SDG-Education 2030 Steering Committee), and reaffirmed the centrality of the Global Education Meetings (to be scheduled in parallel with the UN's HLPF) as the basic structure for providing guidance and reviewing progress in relation to SDG4 (VVAA, 2015b).

Both the WEF 2015 and the FFA came to be perceived as an opportunity to preserve or safeguard the specificity of the education-specific agenda, without compromising on endangering its convergence with the UN agenda. A Member State representative of the EFA Steering Committee thus insisted that the FFA was far from a mere replication of the New York agenda, and that it added key nuance by elaborating on certain key issues "We were trying best to maintain our education-community position. So in the end, what has happened is… you may say they converged, but if you include not only the target-statement but also the notes of explanations, there is some difference, some nuance. Even if the target statement and the overarching statement are exactly the same" (CR5).

Particularly within UNESCO circles, the WEF 2015 and the FFA were viewed as an opportunity to give some form of continuity and to keep the EFA legacy alive, despite the growing awareness that the "EFA brand" was irremissibly lost (and with it, part of the symbolic appeal of the education agenda). The quotes below are indicative of the tension created by the willingness to (symbolically) attach the WEF 2015 and the FFA to the EFA experience, and the need to avoid an image of excessive autonomy vis-à-vis the SDGs. Thus, as the quotes below suggest, the designation of the forum, along with the logo and byname given to SDG4, became an object of struggle. Such branding decisions acquired a high-stakes quality in that they were perceived as central in avoiding an image of agenda-duplicity, but also as an opportunity to signal a sense of continuity to the EFA experience – two objectives that were not always compatible:

---

[119] Note that a provisional version of the text was presented in the context of the WEF 2015. The final wording, however, was made conditional to the final developments of the SDG framework, then being finalized in the context of the UN negotiations, along with the outcomes of the Third International Conference on Financing for Development, scheduled for July 2015 (VVAA, 2015a).

I have been of those in UNESCO that have been fighting for EFA and for keeping this expression [*in reference to EFA*], which for me expresses the right to education as no other formula. But also, of course, it's showing of the continuity of a process that's been lasting for over 25 years. And the World Education Forum, we called it World Education Forum also to show the continuity with Dakar – when others wanted to call it Education Goal Conference or World Conference on Education or whatever, and that was a strong struggle. And we said, "No. We call it World Education Forum because it is the next one after Dakar". And also if you look at the logo and the graphics and the branding, you see that we were still working with a further development of the former EFA logo. And just about a week or two ago, UNESCO came out with a new logo for Education 2030 and SDG 4. And although I was opposed to it at the beginning, I must say it I think it comes out fine. And I think the way the SDG is formulated, the Education for All concept is there. And actually if you take out all the adjectives of the SDG, it still says "for all". And it was something that we also fought for. We fought for every single word and every single word and letter. And the "for all" was something that we would not let go of even if they told us it's too long and it's clear anyway. And we said, "No. No. It has to be written there". And I think that those who wanted to change it, they wanted to show it was the end of everything that UNESCO has been actually quite unjustly accused of […] since we were sort of symbolically holding it [*EFA*], this idea that UNESCO was doing something separate and not aligned, I think we had to overcome that, and maybe letting go of EFA was a way to do that. (UN6)

Some people are positioning… or making it seem like they are two frameworks. And there are not. There is no separate Education 2030 agenda. The FFA is only an implementation strategy. There are no separate SDG4 targets. And it's also about branding, right? So, in the report, there was 2030 in the cover. So what we want to do is put the two things, 2030 and SDG4, in the same frame, to show it's the same thing. (UN26)

It this sense, it is important to bear in mind that the significance of the WEF 2015 and the FFA does not lie exclusively in their content, but also in their potential to frame the new SDG4 agenda as one retaining the most desirable qualities of the EFA experience (i.e., the notion that this is an agenda, crafted by the education community, as opposed to diplomatic circles) while avoiding its pitfalls (its contribution to a dualization of education agendas, its symbolic attachment to UNESCO).

At the same time, the lead-up to the WEF 2015 and the drafting and negotiation of the FFA, became an opportunity to set up a broad, consultative process – one that everyone could perceive as inclusive and participatory, *according to the education-community standards*. Thus, a Drafting Group consisting of the EFA Steering Committee *plus* additional Member States[120] was set up and tasked with the preparation of a preliminary

---

[120] Inputs were thus provided by a broad range of parties, including representatives of Armenia, Benin, Brazil, China, Ecuador, El Salvador, France, India, Japan, Kenya, Norway, Oman, Pakistan, Peru, the Rep. of Korea, Russian, Saudi Arabia, Thailand and Uganda; along with UNICEF, the UNDP, the UNFPA, the UNHCR, UN Women, the World Bank, GPE, OECD, EI, GCE, ANCEFA, ASPBAE and Intel (VVAA, 2015b).

version of the FFA[121]. As noted by a Member State representative of the EFA Steering Committee, broadening the participation was instrumental in order to secure countries' ownership of the FFA:

> A drafting committee is needed to save time for efficiency purpose on each important milestone event […] But for Framework for Action, we thought that because it will be presented to the full participation of member states in conjunction with the general congress of UNESCO… it was known that full member states should participate. And it was better to have broader views reflected in the final adoption. (CR5)

The drafting process is described by a number of interviewees as a highly collaborative and open endeavor. Its participatory and iterative quality came in fact as a surprise (and even as a source of annoyance) to some of the agencies with no prior experience with the EFA community. One of these newcomers was the UNHCR, the lead negotiator of which noted:

> So I represented UNHCR at the drafting committee meetings, and basically, it was, again, many, many, many different drafts, thousands of drafts of this Framework for Action. Really, *we were so sick of it by the end*. I mean, what was surprising to me, every meeting I went to… so the drafting committee process was they would circulate several drafts, and then when they came to a point where they felt the document was ready with the first draft, we had a meeting in Paris, and all of the co-conveners were there, a number of different NGOs and representatives, and then representation of member states, so not all member states were there, but country groups would be represented. So I went to each of those meetings. There were three meetings, and I went to each of them with tons of notes and data. (UN33, emphasis mine)

It is, however, important to bear in mind that, despite the centrality of Member States' input and the contribution from other IOs, UNESCO retained the upper hand in the drafting of the document, particularly during the early stages of the process (see Quote 1 below). It is in fact, worth noting that one of the first drafts of the FFA was co-written in collaboration with a particular federative NGO with a history of cooperation with UNESCO (see Quote 2 below). While this episode was largely the result of personal ties, and ultimately has an anecdotic quality, such dynamics are indicative of the fact that to date, the EFA/UNESCO space remains a much more friendly or beneficial arena for (certain) NGOs – one in which such organizations can resort to informal connections and take advantage of "corridor decisions" which, in the UN sphere, are reserved to another set of NGOs.

> UNESCO, through its secretarial conscience, they really had an influence. So the Assistant Director-General was always there, and also, the director of his unit responsible for EFA was always there… And the Steering Committee chair and the co-vice chairs, these have been consulting actually before and after the community discussions. (CR5)

---

[121] In addition, a first version of the FFA was sent for comments to all Member States before the WEF 2015, as well as before the 38th session of UNESCO's General Conference (UNESCO, 2015b).

I knew [*UNESCO senior official, deleted for anonymity purposes*] from working with [*her/him*] over 15 years on [*deleted for anonymity purposes*]. So I just wrote to [*her/him*] and said, "Do you need a hand in helping to strengthen the draft of that framework for action?" And she said, "Yeah. That would be good. Come over to Paris for a couple of days". Which I did. And ended up sort of helping to rewrite the framework in UNESCO, and then sort of sharing it with… basically, sort of getting it through. I'm trying to remember who the people were now, the Assistant Director-General for education, and others, who seemed quite happy with the revised draft. And quite a few of the things that we were able to put in at that point stayed in right the way through to the approval of the Framework for Action, including much stronger language on financing, I think. And sort of tighter definitions of roles and responsibilities, centrality of government, a stronger rights-based framework than was otherwise there, and so on. (CSNGO3)

At the same time, and given the need to ensure a real sense of country ownership, the later stages of the drafting process were very much steered towards securing the Member States' buy-in of the FFA. In addition, as a result of the willingness to prove that the Education 2030 agenda was a collective endeavor (as opposed to EFA, eventually perceived as UNESCO's agenda), the FFA process was particularly permeable to the lobbying efforts of a wide array of organizations (which frequently pushed for the advancement of particularistic agendas) and conceded substantial *veto* power to Member States. In order to avoid alienating any constituency or country, consensus-making practices typically relied on an aggregation of preferences. Any proposal was likely to be retained unless proved unacceptable to a particular party (and if this were the case, the language would be tweaked until the issue did not represent a deal-breaker to any constituency[122]). Conversely, political trade-offs and potentially divisive issues were deliberately avoided[123]. Such dynamics might explain the expansive or "Christmas-tree" nature of the final FFA:

> You know, every subsector does its lobbying. So every subsector lobbied […] So, there were not tensions. No, I don't think there were any particular tensions. Because ultimately SDG4 it's about any education system, from early childhood to adult education. It's

---

[122] An example of these *word-smithing dynamics* can be seen in the language around financing. Different interviewees noted that references to the percentage of GDP and public expenditure allocated to education were a recurrent object of negotiations, with the language being tweaked in order to make the final working acceptable to all parties. At the same time, consensus in this particular area also relied on *forum-shifting dynamics*, with those leading the FFA drafting efforts insisting upon the fact that questions relative to domestic financing and ODA were not the responsibility of the SDGs but that of the AAAA negotiated in parallel and formally independently of the SDG deliberations (see section 5.1.2 above).

[123] The deliberate effort to circumvent those debates, perceived to lead to ideological polarization, can be clearly observed in the FFA language around the role of the private sector. One of the UNICEF negotiators noted that, "Arguments on private involvement seem to be black-and-white issues, it tends to become ideological […] Because the positions of the different partners were so apart, it would have been *impossible* to come to what I would say an appropriate compromise" (UN30). Such conflict-avoidance efforts were also recollected by one of the Member State representatives, leading the drafting efforts, who observed, "I think if you read the Framework for Action it is not very specific when it comes to delivery, it is saying who is responsible and who should be the main financing part. But how it should be delivered… we have not been that specific about the delivery part, *and that being deliberate*" (CR2, emphasis mine).

everything [...]. Which is why, I guess, this is a consensus. I don't think there were necessarily tensions that needed to be overcome. (UN20)

Everyone met at this meeting with their own agenda of what they thought should really be in the framework, but I surprised to find that in general, developing the Framework the Action was convincing the member states. They were the ones who really had the power to say, "We don't want this paragraph here", or, "We don't want this wording like this" (UN33)

The consensual quality of the FFA, along with the widespread perception of the debate as a low-stakes process (for the final list of goals remained the responsibility of the IGN), might explain the fact, noted by certain interviewees, that the WEF 2015 was something of an anticlimactic moment – one lacking the tension or unpredictability that had characterized the Dakar and the Jomtien meetings held in 1990 and 2000. This was reinforced by the general perception that the WEF was a heavily-scripted, stage-managed meeting. Reflected on such dynamics, one participant stated explicitly, "I feel a little bit that UNESCO's a little bit of an old-school agency in that it has these big meetings. Very bureaucratic-driven, it's very bureaucratic. It has these huge meetings in which the member states come, blah, blah, blah. It's a little bit old-fashioned" (UN33). At the same time, a range of interviewees shared the notion that such meetings, even if performative in nature, were instrumental in providing further legitimacy to the new education agenda – although, in practice, the targets were being discussed elsewhere:

You know, it's a result of a consensus, but honestly I was at the World Education Forum, and the adoption of the Incheon Declaration was kind of without discussion… I mean it was quite surprising who has followed the global agenda for some time, it was kind of a let-down, a disappointment for critical researchers who were looking to see what would be the debates and the tensions [*laughter*]. You know, there was very little debate, I mean, there were one or two issues, on one or two words, and basically the adoption. Applause, and adoption and whatever… so there wasn't much tension. (UN32)

So exactly what is left for Incheon? If the goals have been decided, the targets had been decided, and the framework for action is not being discussed, what exactly is Inchon about? I: It was like, a signification… it was good to have had it, in terms of people who attended. I've talked to people who were there afterwards, they say it was good to meet all these people from other countries, and interact and learn and share, and find out how things are going in other places, and so on. But it wasn't really a significant decision-making forum in that sense. (CSNGO10)

You know, it all was more or less decided. But in having the World Education Forum, having so many ministries… The momentum was… Whatever that came in September, it was given further legitimacy. (UN1)

In this sense, the symbolic potential of the WEF 2015 and the subsequent adoption of the FFA, can hardly be overstated. Such episodes proved powerful opportunities for the global education community to regain a sense of unity, consensus and cohesiveness. In particular, the revitalization of the inter-agency coordination mechanisms in the lead-up to the WEF 2015, and the drafting of the FFA, symbolized a turning point – for, as previously noted, such processes sparked a considerable level interest among global

education agencies (a behavior that contrasted greatly with their disengagement, during the tail-end of the EFA). Within certain UNESCO education circles, this was perceived as a marker of success and revitalization – the proof that UNESCO continued to retain a certain capacity to lead within the UN system while being perceived as an honest broker able to bring the global education community together and build consensus in a pluralistic environment. The sense of satisfaction is captured in the words of a UNESCO negotiator (see Quote 1 below) but also of a CSO organizer, who had similarly high expectations on the potential of the Steering Committee, and the comparative advantage that the (post)EFA architecture afforded to the education sector, in contrast with other SDG areas (see Quote 2 below):

> When we started the consultation process and the EFA Steering Committee, some of the other agencies were just consistently absent and not responding and just showing their complete disengagement with the process. And then suddenly 2015 came around and not only were they all there and are now, but also new agencies came. And the fact that ILO and UN Women and UNHCR were coming and asking us, "Could we please also be, what do we have to do to be a co-convener?" […] I think that should also go actually on the success story of UNESCO, this capacity at a very difficult time and with a lot of competition for attention, resources, goals of the agencies within the UN system, how we managed to mobilize new partners […] I'm actually very proud of how UNESCO came out of the process and of the result because I think it has… I think that UNESCO has managed not only to preserve but to considerably reinforce its image as an honest broker, as an organization that is really working for an ideal, the right to education, and that is not pursuing any of its own agenda. (UN31)

> In some ways thanks to the EFA goals, UNESCO has been ahead of many other agencies in the sense that they actually had a structure already, the EFA starring committee, the fact that they had a Framework for Action, I mean I don't think any other goal has that kind of a operationalizing attempt in place yet […] And I think what you see now is quite a broad sign up from different UN agencies to the Framework for Action, and they will also be represented in the Steering Committee or the kind of Education body that will oversee the implementation of the Framework for Action and the SDG4, I guess. (CSNGO1)

At the same time, the engagement of most education agencies, particularly in the context of the SDG4/Education 2030 Steering Committee, has progressively revealed itself as somewhat superficial and short-lived. While such developments are beyond the scope of this dissertation, it is important to bear in mind that there are signs, that the SDG4/Education 2030 architecture could soon regress to the EFA dynamics. Thus, in a follow-up interview conducted in 2019, a UNESCO senior official regretted the progressive disengagement of certain key IOs from the Steering Committee, as well as the lack of interest generated by SDG4, *even within UNESCO* – a dynamic was all too reminiscent of the withering and decline, experienced by the EFA architecture in the aftermath of Dakar (cf. Chapter 2). At the same time, the interviewee recalled that only the 2018 Global Education Meeting (held in Brussels and bringing together some 400 participants, including a number of Ministers of Education and Development and delegations from more than 60 countries – cf. UNESCO, 2018) had managed to spark a certain interest among the rest of the SDG4 partners, and within UNESCO's top-

management and high-ranking administrators. Such views are echoed by a recent evaluation of UNESCO's education sector, which recalled that the momentum built in the lead-up to SDG4/Education 2030 had been lost, and noted that "UNESCO had not sufficiently planned for the transition from agenda- setting to coordinating implementation and monitoring progress […] Key external stakeholders consider that the global and regional architecture for supporting implementation of SDG4-Education 2030 is not functioning effectively" (UNESCO Internal Oversight Service, 2019, p. vi-vii). Overall, such dynamics suggest that, while UNESCO retains significant convening power, its role as facilitator and as the focal point for global education debates is far more uncertain.

> So in the first year, there was a lot of interesting participation at the high level in the Steering Committee, and then they took over the SDGs within their own organization. But from the second year onwards, from UNICEF, from the World Bank, the other UN agencies, UNFPA, UNDP, there was a kind of disengagement in the way they attended Steering Committee meetings… Sometimes not even a senior-level official would attend, and there was no kind of intention to actually contribute to the decision-making […] And then [*In reference to the 2018 GEM*] it really energized people. It did give members some sense that UNESCO is doing something. I felt very gratified. People appreciated UNESCO's role and at least initially, we saw a lot of interest from partners, the World Bank, UNICEF, a lot of others, even a lot of the countries. (UN30)

> One of the failings is that the Education 2030 division is just under-resourced. There is a lot of structural discussion. The only area we advanced on was on the indicators. And largely because the UIS set up a separate structure to bring people together […] The main challenge comes back to the fact that UNESCO does not have enough of resources to drive the agenda. I mean, I think the internal shifting of emphasis in the last year towards intelligence, the Futures of Education…. In a way it undermined and moved away from the SDGs. In conceptualizing the Futures of Education world, there was not enough discussion on what does it mean for the SDGs. There was no discussion actually […] The value of the SDGs is only recognized when we have a global meeting. When we had the meeting in Brussels [*name of a high-level UNESCO manager, deleted for anonymity purposes*] recognized that it was a big success. All the ministers came and everybody was [*inaudible – there?*]. But the month after, there was not much interest in how do we take that forward. It was left to a small team to kind of work, with the task to see how we could move things forward. (UN30)

### 5.2.7. From the TAG to the TCG: removing politics from science?

As discussed in the prior sections, the negotiation of SDG4 proved a challenging process, in that it required finding consensus on which deliberation *fora* constituted legitimate decision-making structures. After the adoption of the 2030 Agenda for Sustainable Development in September 2015, the debate was partially settled – with the EFA architecture being organically merged with the broader UN process, and the education community preserving a sense of autonomy through the coordination and monitoring mechanisms established by the FFA. However, conflicts over the decision-making architecture and the division of labor resurfaced in relation to another branch of the SDG4 negotiation – namely, the development of the global and thematic indicator frameworks.

Thus, the transformation of targets into quantitative indicators proved a divisive issue which, once again, pitted the education-specific *fora* against the UN negotiation chambers. On this occasion, matters were further complicated by the attempt to establish a neat division of labor between the political and technical areas of the debate – a division that reignited fears over the loss of centrality of the post-EFA architecture. The following two sections grapple with such dynamics by examining the debates, triggered by the very *conceptualization* of the SDG4 indicators, i.e., the process through which the SDG4 indicators were named and formulated (cf. Chapter 3).

For contextualization purposes, it is important to bear in mind that, in the area of education, efforts to devise an indicator framework did not start with the UN General Assembly's decision to delegate such work to the UNSC (cf. section 5.1.3 above), but dated back to the mid-2014 and, more specifically, to the establishment of an inter-agency, ad-hoc platform known as the Technical Advisory Group (TAG). The TAG was established by UNESCO and recruited experts from UNESCO itself, but also from the GMR, the OECD, UNICEF and the World Bank. Chaired by the UIS, it was tasked with the preparation of a series of recommendations for what came to be known as "post-2015 indicators" – a denomination reminiscent of the fact that, at the time, the relationship between the SDGs and the post-EFA process was still far from clear.

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. Thus, on the basis of the TAG's first proposals, participants were invited to evaluate the alignment of the proposed indicators regarding the (expected) targets, to rank the proposed indicators (identifying the most relevant for each target) and to propose alternative indicators (Motivans, 2015).

It is important to bear in mind that since the indicator process started before the end of the negotiation of the targets, certain TAG members had the expectation that the "technical debate" would shape the political deliberation. The excerpts below, corresponding to an OECD- and a UIS adviser, capture such line of reasoning:

> So the technical dimension is grounded in realities, it's grounded in what can be measured now, what could be measured 5 years from now, what could be measured 15 years from now. And then the political dimension is "What are the aims? What are the goals we are setting ourselves as humanity? Don't bother me with whether it can be measured or not! I'm setting and aspirational target". So that's what you're dealing with. And then at the political level also, you have very powerful lobby groups who refuse to let go of certain areas. Even though something can't be measured, they insist that it has to be included because this is politically so important. (LASS3)

> I have to say that the final list didn't quite succeed on this... but what at the beginning of the TAG we tried to do was to propose indicators that we though were feasible. So they either existed, and had a reasonable coverage, or we thought they could be extended

relatively straightforwardly to cover, you know, a sufficient number of countries. And to be available with a reasonable frequency, but not necessarily annual. But when the list was finalized, some other indicators appeared in the list, let's say [*laughter*]. So not all indicators really passed that initial test that the TAG had set itself. (UN11)

A final proposal incorporating the input from the consultation with the education community was presented in one of the WEF 2015 parallel sessions (*Framing and developing indicators to measure progress for the 2030 education targets*). It was only after that meeting than the political implications of the SDG4 quantification process started to be more widely perceived, or at least more explicitly acknowledged. On the basis of this, an "extended TAG" (ETAG) was created. The ETAG incorporated representatives from Member States[124], along with experts from civil society organizations. The ETAG was tasked with the refinement of the list of indicators to be included in the Education 2030 FFA (UNESCO, 2015c), and to this purpose it celebrated two in-person meetings in July and September 2015. It also benefitted from the input of yet another consultative process undertaken by focal points (Montoya, 2015). This work culminated in a second indicator framework, which, for reasons discussed below, was included in the FFA, *only as a proposal*.

With the new Education 2030 agenda formally adopted, the ETAG mutated again, giving rise to the Technical Cooperation Group (TCG), which remains operative to date. The TCG was thus established in 2016, and was conceived as a space for discussion, as well as a technical platform to support the UIS in the implementation of the thematic indicator framework. This group was also expected to assist the IAEG, the SDG4 Steering Committee, regional bodies and countries in their data-collection and reporting efforts. Chaired by the UIS and the UNESCO Education Sector's Division for Education 2030 Support and Coordination, the 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 organizations and the Co-Chair of the Education 2030 Steering Committee (see UIS, 2017a).

The transition from the ETAG to the TCG was not without problems. Some original members of the TAG perceived the TCG as a marker of the increasingly politicized nature of the indicators debate. At the same time, certain countries represented in the TCG perceived that their input had not been sufficiently taken into consideration – but simply used for rubber-stamping purposes. A Member State representative involved in the TCG over a long period of time elaborated on such tensions in the excerpt below:

> Sometimes, you make a suggestion as a country, and then you kind of get the impression that there's a lack of transparency. It' a bit of a frustrating process […] So I guess it's more a concern with the process. It's a consultation and we can effect change. And they seem willing to listen. But ultimately, sometimes we're very constrained. There's such a speed that goes on and kind of a pressure to make a decision. And it's very difficult, I

---

[124] More specifically, two experts from each of five regions, from Member States that were also members of the IAEG-SDG or HLG.

think, for somebody to provide true input when they're struggling just to find their words. (CR4)

Thus, despite the line of continuity between the TAG, the ETAG and the TCG, and while their mandate remained similarly constrained to technical questions, it is important to bear in mind that such mutations had an important impact on the internal dynamics of the group. Even if the transition to the ETAG was perceived by some as cosmetic in nature, it entailed an important change in the decision-making dynamics and the very atmosphere of the group. This is so as one of the most remarkable aspects of the original TAG had been the virtual absence of conflict or disputes, despite the high stakes associated with their mandate. Different interviewees partaking in this first stage of the TAG recalled a sense of collegiality and concerted effort. They described the process as one in which the representatives of different organizations put technical considerations before political or institutional interests – even if, as reflected in the words of one of the statistical experts involved, the TAG members were not fully oblivious to the existence of organizational priorities and interests:

> From the TAG angle we were all pretty active in the beginning. So there was OECD, the GMR as it was then, UNICEF, the World Bank and the UIS […] We were all pretty active. You know, some of us worked on indicators, some of us worked on the text of the document about the thematic framework… and we had quite in person, like detailed debates on, you know, whether this indicator really works in practice or not… So, it was very collaborative work. Obviously each individual in the group had different areas of expertise, or different interests, either coming from themselves, or perhaps from their organizations. So it wasn't that we were all working in every single target at the same time. So obviously people like the OECD was represented by a PISA person, so they had a stronger interest in the learning outcomes indicators… but not exclusively so. (UN11)

There is thus a general agreement that, despite the variety of institutional priorities, the TAG-led process did run smoothly. This cannot be separated from the fact that the TAG comprised a limited number of members selected on the basis of their technical expertise (i.e., statistical competence) and their privileged understanding of international education statistics, and who shared a similar professional background. Thus, the members of the first TAG perceived themselves as peers in the professional sense of the word, and were keen to collaborate, under the assumption that their work would be evaluated on technical grounds. Similarly, they perceived themselves as a cohesive group working against a common threat – an excessively ambitious (and consequently immeasurable) list of targets. The following words of a World Bank senior official captures this sense of separation between political and technical work, and the perception that a consensus is always workable – provided the conversation is held in technical terms:

> We are a technical group, and all ideas are open for discussion. So… you know, all the institutions have their own goals, and the Bank has its own goals […] But at our level, in the technical group, I think we are moving in the right direction. But then again, the only thing that we can do is to make recommendations. It is the higher level that really determines the whole process. At the higher level, it's what makes the difference. You know, there are groups that are fighting on the whole thing. Cause we can only give

advice… At our level there is little politics, but when it goes up, it goes to the political level. Because we are looking *scientifically*, so this is what we find. And I don't think there is an issue that cannot be solved technically. (DON2, emphasis mine)

The sense of like-mindedness and concord is particularly apparent in contrast with the dynamics that prevailed in the ETAG and, subsequently, the TCG. As recalled by different interviewees, the incorporation of country and CSO representatives entailed a dramatic change in the tone of the conversation, which simultaneously became more permeable to political and ideological considerations, and less driven by the (self-attributed) sense of realism and pragmatism that the initial TAG team had consciously cultivated. In fact, various interviewees coincided in noting that the shortcomings of the final list of thematic indicators were, at least to some extent, the product of the democratic character of the process. While there was a general acknowledgement of the importance of countries' input, some perceived that the democratic character of the process could have been better served by other mechanisms (most notably, consultations) rather than by bringing countries to the table, at least in the initial stages. When probed about the pros and cons of the ETAG structure, a UIS analyst noted:

I think we made a conscious decision at the beginning, not to bring in countries at the start. Because you needed to have a knowledge and an expertise on international indicators. And most countries don't have that […] And I think it might have made it more difficult to reach a consensus, because they would have been so many ideas coming in all sorts of… interesting and unusual indicators… but maybe not be particularly practical. And I guess that we are doing this [*extending the TAG membership*], because the ideal now is that, as much as is possible, these things should be country-led […] But what we tried to do, even when we didn't have countries there, we already had a global consultation on the original framework. So we tried to bring in ideas, but obviously we had to come up with the framework to start, with people commenting on we had purposed (UN19)

However, according to certain negotiators involved in the indicator development process, the element perceived as more obtrusive was the *veto* power enjoyed initially by the SDG4 Steering Committee. As recalled by various interviewees, certain indicator proposals around which there was a reasonable consensus within the ETAG or the TCG, were recurrently blocked or called into question by the Steering Committee, sometimes at the request of some of the (dissident) members of the ETAG. In view of the impossibility to "tame" the Steering Committee, certain agencies resorted to a circumventing strategy which devolved the decision-making power to the technical body – namely, the TCG. The reasoning behind this deliberate shift in the *locus* of power is captured in the words of two of the negotiators involved in the indicator development process, who noted that:

Because some of those things [*controversies around certain indicators*], it made very difficult to get the Steering Committee to approve the list. So at that time we thought they would need to be the approvers, and it became clear that this was never going to happen. For whatever reason, they could not reach a consensus. So, last year we sort of put the responsibility back to the technical level, to get to make the approving. And we were able to reach that consensus in Madrid [*TCG meeting in 2017*] […] I think you saw some

countries had reservations on certain indicators that are up for development… And I think that's always going to be the case. You'll never going to get full agreement. But I think because the global indicators were accepted by the Stats Commission, I think this helped the TCG to be able to say "Let us be the ones that approve", and then it went to the Education 2030 Committee just for them to endorse, or like note that this agreement had been reached. (UN11).

When the UN piece went on its way, and the UN technical agencies, including the World Bank, had a certain amount of inroads into supporting that, there was reasonable coherence across UIS, UNICEF and the World Bank, in thinking that we should be pushing forward on the metrics. However, there was also the other consultative group… I'm thinking particularly of the SDG4 Steering Committee, which ultimately actually has no power. I mean, that's one of the biggest issues with the Steering Committee, it steals people time. It helped with the Agenda for Action [*FFA*] and these pieces, but it didn't have much power in the technical work. And yet it created its own team to do work on that, and in that team there were people who were very anti the early grade reading metrics, in particular. and that kept coming up […] The pushback was coming particularly from civil society and states like France, on behalf of Group 1, even if this group represented also the UK, which was very supportive of the early grade reading. So you had the situation in which the Steering Committee did not have that much power into this, but having important things to say, they overstepped… And people in the Steering Committee, they were not representative of their groups, which undermined the whole Steering Committee. So that was probably the most difficult part. And in some way the technical agencies we kind of ducked and bypassed the fight because… it was going to get through! (UN25)

Ultimately, such dynamics are indicative of a certain resurfacing of the old tensions around the centrality and decision-making power, merited by the EFA architecture, even if, on this occasion, such frictions were presented as a result of an incompatibility between an expert, apolitical approach, and the more ideological perspectives brought by Member States and CSO representatives.

### 5.2.8. Beyond the education realm: the negotiation of the Global Indicator Framework

As previously noted, in December 2014, the identification of the global indicators became the responsibility of the UNSC which subsequently established an Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs), responsible for the development of a global indicator framework. The rapport between this body and the structures put in place by the global education community soon proved problematic – for it revealed that the autonomy and authority of the post-EFA structures vis-à-vis the UN architecture were severely limited.

Thus, while the TAG (and its successors) retained responsibility for the thematic indicator included in the FFA, they lost the capacity to directly shape the global indicator framework (i.e., the set of globally comparable indicators to be endorsed by the UN General Assembly). This was all the more relevant given that that the global framework was expected to have a direct impact on the FFA – the global indicators were requested to be incorporated into sectorial frameworks. The prevalence of the global framework

over the thematic framework was made clear by the UIS Director in her presentation of the ETAG Report to the EFA Drafting Group in October 2015, when she noted, "The final list [*of thematic indicators*] will be defined when the global indicators are adopted. The global indicators will make it to the list of the thematic indicators" (Montoya, 2015, u/n).

To be sure, the influence of the TAG during the initial stages of the global indicator framework deliberations can hardly be overstated. Despite the short-lived nature of the group, the preparatory work it conducted significantly shaped the global indicator framework. Thus, the first proposal of the TAG fed directly into the preliminary list of indicators compiled by the UNSC by January 2015 – a list that would eventually form the basis of all negotiations relative to the SDGs' measurement efforts. It is important to bear in mind that the list was prepared on the basis of the submissions of experts from international agencies, organized in goal-specific clusters that mirrored the structure of the TST for the OWG on SDGs (UNECE, 2015). Thus, while the UNSC never approached the TAG, the work of the TAG constituted the basis on which UNESCO and UNICEF, as education co-leads in the UN Technical Support Team, provided recommendations to the UNSC (UNESCO/UIS, 2015).

The role of the TAG is also central in explaining the high levels of inter-organizational consensus that characterized the process of refinement of this initial proposal for the global indicator. The existence of a certain concord between the main education agencies involved in the process became apparent in the consultation run by the UNSC until May 2015 with the explicit objective of narrowing down the first proposal. While the initial list contained a maximum of two indicator proposals per target, in the context of this consultation, specialized agencies were asked to select a "priority indicator" for each target. Quite significantly, UNESCO and UNICEF exhibited a significant degree of consensus and coincidence in their preferences, generally identifying the same "priority indicators". Later on, in the first open consultation, conducted by the IAEG, education agencies exhibited a similar degree of consensus – UNESCO, UNICEF, OECD and the World Bank consistently selected and supported the same "priority indicators" (see IAEG, 2015[125]).

However, it was also in the context of this refining exercise that a breach between the TAG proposal and the one embraced by the IAEG started to emerge. The discrepancy between the indicators favored by the global and the thematic frameworks became particularly obvious in relation to Target 4.1, relative to learning outcomes. As synthesized in Table 5.2 below, the proposals advanced by the TAG and the ETAG (in May and October 2015, respectively) only took into consideration the learning levels of children at the end of primary and secondary education. In concordance, the indicator that

---

[125] Of particular interest are the Contri*bution of UN Statistical System organizations to the work of the IAEG* as of 5th September 2015, as well as the *List of Indicator Proposals*, made available on 11th August 2015.

appeared in the *preliminary* version of the FFA (as published in November 2015) focused on the percentage of children at the end of primary and at end of lower secondary achieving at least a minimum proficiency level in reading and mathematics. Conversely, a slightly different indicator was making process in the IAEG-led workstream. Thus, the proposal advanced by the IAEG also contained an indicator relative to the learning levels of early graders (Grades 2 and 3). Even if this difference might be perceived as minor, the inclusion of the early grades was perceived as a red line on the part of certain CSOs and countries sitting in the SDG4 Steering Committee.

The incorporation of the reference to early grades into the global indicator framework is generally considered to be the result of the input of certain UN Member States. A number of interviewees coincidentally noted that the lobbying efforts and pressures on the part of the US were central in the inclusion of such reference – a move that some attributed to the organic ties of the US Agency for International Development (USAID) with a particular data producer whose activity focused on the assessment of early graders:

> These are tools developed by USAID, there is clear interest in scaling those up. They were more quietly part of the agenda through negations, but maybe about a year ago it seemed that the US really started to push hard on that, started making calls, organizing the community, wanting to push this forward, and really ensure that you had an indicator on early grade reading. (CSNGO11)

> I think, honestly, that the reason that 4.1.1 was framed as the measurement in second or third grade was the pressure from [*name of the organization, deleted for anonymity purposes*] to have a measure that would stimulate the demand for their tool. (PRI8)

> I do think that there was the potential for USAID to not get behind the SDG for education, because during the process there was a proposal to only start measuring around the end of primary and the idea that the SDGs shouldn't focus on measuring any kind of learning in the early grades of primary. And that was something I think that if the target and indicators had not included the Grades 2 and 3 component, where countries are required to report out on reading and math, I think then it would've been hard for USAID to get behind this SDG agenda because their own strategy is so much based around early-grade reading. (DON1)

However, while the inclusion of the early grades in indicator 4.1 is generally considered the result of the preferences expressed by Member States, a review of the proposals and comments, submitted by different education agencies suggests that such changes were also informed by the recommendations advanced by UNICEF (and, to lesser extent, the World Bank). As reflected in Table 5.2. below, certain education agencies were supportive of an early grade indicator in the context of the first IAEG meeting and, subsequently, reaffirmed such support in the context of the joint proposal made by the Chief Statisticians of the UN. In this sense, it is interesting to note that, while during such processes UNESCO stuck to the consensus reached in the context of the TAG, other agencies represented in the TAG *went off script*. More specifically, whereas UNESCO consistently suggested a subset of thematic indicators agreed in the WEF 2015 to be used for global monitoring purposes, this was not the case of other agencies. In other words,

the existence of two parallel workstreams (one led by the TAG, and one led by the IAEG) was strategically used by certain international organizations as a means of advancing their own preferences, even if these did not generate consensus within the (post) EFA community.

**Table 5.2. Indicator proposals relative to target 4.1 – IAEG negotiations**

| Forum or consultation | Indicator proposal |
|---|---|
| TAG proposal – WEF 2015 (May 2015) | Percentage of children who achieve minimum proficiency standards in reading/mathematics at end of: (i) primary and (ii) lower secondary school. |
| Proposal sent by UN agencies to UNSD – May 2015 | UNESCO proposal: [Percentage of children/young people at the end of each level of education achieving at least a minimum proficiency level in (a) reading and (b) mathematics.]<br><br>UNICEF proposal: [Percentage of children who achieve minimum proficiency standards in reading and mathematics at end of: (i) Grade 2; (ii) primary; and (iii) lower secondary.] UNICEF suggest the inclusion of "grade 2" as a critical stage for monitoring children's learning.<br><br>WB proposal: the indicator requires the development of a global metric for each subject as a reference point to which different assessments (national, regional and international) can be anchored. Assessments at other levels (e.g. Grade 2) could be considered. |
| Comments by UN agencies sent in the context of the IAEG Open Consultation | Joint proposal of the UN Statistical System, supported by UNESCO-UIS, UNICEF, OECD and the World Bank, *i.a.*<br><br>Percentage of children/young people (i) in Grade 2/3, (ii) at the end of primary and (iii) at the end of lower secondary achieving at least a minimum proficiency level in (a) reading and (b) mathematics. |
| Inputs from agencies and other entities on indicator proposals and metadata in the context of the 1st IAEG meeting (as of 15 June 2015) | UNESCO-UIS proposal: Percentage of children who achieve minimum proficiency standards in reading and mathematics at end of: (i) primary (ii) lower secondary.<br><br>UNICEF proposal: Percentage of children who achieve minimum proficiency standards in reading and mathematics at end of: (i) Grade 2; (ii) primary; and (iii) lower secondary. UNICEF suggest the inclusion of "grade 2" as a critical stage for monitoring children's learning. |
| Input from the Statements by UN Statistical System Organizations to the 2nd IAEG meeting (26-28 October 2015) | Modify proposed indicator to include assessment at the early grades of primary and clarify the levels at which learning assessments should be made:<br><br>Percentage of children/young people (i) in Grade 2/3, (ii) at the end of primary and (iii) at the end of lower secondary achieving at least a minimum proficiency level in (a) reading and (b) mathematics). |
| Consolidated target as of the 3rd IAEG meeting (30 March – 1 April, 2016). | Proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex. |

| Indicator approved in the 47th session of the UNSC | Percentage of children/young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics. |
|---|---|

Source: Author's elaboration.

The divergence in the proposals advanced by the TAG and the IAEG was eventually settled in favor of the latter – as, according to the hierarchy of monitoring levels established at the start of the process, the thematic indicators needed to incorporate the global ones. In practice, thus, the priorities established the TAG and the ETAG were superseded by the ones established in the context of the UN process. This meant that the list of indicators included in the FFA was eventually incorporated some indicators that have not been agreed upon by the SDG4 Steering Group or in the context of the WEF 2015.

The unrest brought about by such dynamics crystallized in the First Meeting of the SDG-Education 2030 Steering Committee, held in May 2016. In the context of this meeting, representatives of the civil society expressed their concern regarding the inclusion of Grades 2 and 3 in the indicator framework, noting that this diverged from the consensus reached in October 2015 – when the EFA Steering Committee had agreed on *not* incorporating a reference to early grades in Target 4.1.1 (a consensus reflected in the provisional FFA approved in November 2015). In response to such concerns, an agreement was reached for the Steering Committee Co-Chairs to send a letter to ECOSOC reasserting the position of the committee (UNESCO, 2016a; see also ASPBAE, 2016; or UNESCO, 2016b).

The centrality of the IAEG vis-à-vis the TAG (and its successors) and the SDG4 Steering Committee thus appears to be at the root of the discontent expressed by some constituencies in relation to the construction of the SDG4 indicator framework. Again, tensions around the post-2015 debate do not result exclusively from its very *outcomes*, but from the mechanics and the architecture of the deliberation process. Thus, a number of interviewees sitting in the EFA and the SDG4 Steering Committees voiced strong criticisms regarding the power exerted by Member States in the context of the IAEG, arguing that such spaces lacked the transparency, inclusivity and democratic *ethos* that characterizes the post-EFA architecture. Similarly, various interviewees were critical of the attempts to "depoliticize" the construction of the indicator framework. It was noted that demarcation between political and technical labor was an artificial division – and one likely to lead to a reinterpretation of the SDGs, thus torpedoing or undermining the consensus built during the negotiation of the goals and targets. Such perceptions are most clearly articulated in the excerpts below:

> What the member states did when delegating the global level indicators to the Statistical Commission was actually really unfair, because they gave a task to the statisticians that they couldn't do well […] Because, they would have needed somebody, some political guidance to tell them on what aspect of every target to focus. But the only political guidance they got is "One global indicator per target, please". And what has happened is

that they've allowed the availability of data to determine what they measure under each and every target, so basically we're re-writing history completely, and we are re-negotiating the targets without having negotiations but this kind of pretense of the technical absolutely neutral process. So the UN agencies have had a lot of power and so have the twenty countries that are part of the expert group of the Statistical Commission that has been working on this. And then, to make matters worse, they've skipped a couple of steps in the process, which means that they've made a lot of decisions indirectly. (CSNGO1)

And this [*the establishment of the IAEG*] has resulted in discussions led by technical experts who are indicator experts. And this raises also additional concern. On one hand, it is very true that Steering Committee members who represents each constituency and gather to discuss are not necessarily indicator experts, so it becomes very difficult for this Steering Committee to decide on indicators. But experts on indicators are not necessarily aware of the broader concerns of education. They are indicator technical experts. So that has, I think, provided some source of concern for the broader education stakeholders. (CR5)


## 5.3. An unfinished learning turn? SDG4 as a site of normative struggle

While the architecture of the post-2015 education debate represented a recurrent point of contention, it was not the only area in which a consensus needed to be found between different and sometimes competing interests. Predictably, the negotiation of SDG4 also required finding a compromise between parties with different interests and expectations on to the very *content* of the new education agenda, i.e., the principles and priorities that should underpin it.

In this regard, one of the most persistent areas of contention was the status of learning in the global education agenda. Thus, certain constituencies seized the post-2015 process as an opportunity to frame the *improvement of learning outcomes*, as the cornerstone of the new education agenda. Such actors had long been seeking to place learning at the center of development efforts, and found in the post-2015 debate, an opportunity to imbue the improvement of learning outcomes with the norm-like status once exclusively enjoyed by education access and the universalization of primary education. However, other constituencies identified a number of risks in this normative shift, and actively resisted the attempts to frame the improvement of learning outcomes as an overriding priority. In this sense, the negotiation of SDG4 operated as a site of normative struggle that rendered visible the existence of different assumptions and understandings regarding the ultimate purpose of education.

This section aims at tracing and explaining the intersection of the post-2015 debate and this the broader process of normative change. To this end, the section first explores the origins of the learning turn, understood as the intent to consolidate the improvement of learning outcomes as the prime concern of the education-for-development field. The section follows with an overview of the multiple initiatives that have attempted to advance this agenda in the context of the post-2015 debate, as well as the efforts made by

certain actors to keep such ambitions at bay. The section continues with a discussion of the (fragile) compromise reached between the two camps at the end of the process, and finishes with an overview of recent initiatives through which the learning agenda has continued to advance – despite its limited success in the SDG4 context.


*5.3.1. Laying the ground: human capital theory revisited*

Efforts to give greater prominence to *learning* as opposed to *schooling* date back to the early 2000s. The ideational roots of the learning turn are in fact located within an organization that, as can be inferred from the prior section, had only a limited presence in the negotiation of SDG4 – the World Bank. Along with some affine researchers and commentators, the World Bank played an instrumental role in theorizing the need for a normative change within education development circles, and in disseminating the view that the emphasis on access brought about by the MDGs was ill-founded and misleading. Thus, a series of studies, conducted by well-known economists connected to the World Bank sphere, introduced a crucial tweak in the human capital theory – one that would eventually translate into a reorientation of the work of a wide range of development partners. Two seminal research papers found that the knowledge and skills acquired in primary education, rather than the number of completed schooling years, were the element making the difference in terms of national economic growth and individual mobility (see Glewwe, 2002; and Hanushek and Kimko, 2000). Subsequently, a paper, authored by Eric Hanushek and Ludger Woessman (2008) insisted similarly on the argument that the level of skills, as measured by performance in student assessments, was a better predictor of economic growth than schooling. The decoupling of schooling and learning was indeed a disruptive move for development circles had long assumed a tacit connection between schooling years and educational progress. The significance of the new line of reasoning inaugurated by these authors is captured in the words of an officer of a bilateral organization, who noted:

> And suddenly you have economists like Hanushek saying that learning is important for economic growth, and peace and stability […] Because in the beginning in the 60s and 70s, the World Bank did all of that human capital research which equate all of that economic growth with access to schooling. And then Hanushek came in and all changed… Because people thought that access meant learning. No-one imagines that all these kids would go to school and not learn anything! No-one imagined that you would train teachers and buy books and that they would learn nothing! It wasn't until we started testing and the civil society organizations started testing, that we saw that that was... So then Hanushek came and said "It's not schooling, it's actually learning" (DON4)

The publication, in 2006, of an evaluation of the World Bank's work in the area of education represented a milestone in the institutionalization of such ideas within the development realm. Informed by the findings of the aforementioned research, the report was explicitly designed to orient the organizational priorities of the institution. In this sense, it represented a turning point in the primacy conceded to the learning outcomes theme vis-à-vis universal enrolment and completion. Conducted by the Independent

Evaluation Group, and entitled *From schooling to learning*, the report found that, while the objective of access expansion had been successfully met, insufficient emphasis had been placed on the improvement of learning outcomes (IEG-WB, 2006). The recommendations given on the basis of such findings were unambiguous on the need to transcend the MDG focus on access and completion, and to reorient the World Bank and the FTI activity in the area of primary education, in order to ensure greater priority for learning outcomes.

> Primary education efforts need to focus on improving learning outcomes, particularly among the poor and other disadvantaged children. The MDG push for universal primary completion, while a valuable intermediate goal, will not ensure that children achieve the basic literacy and numeracy that are essential to poverty reduction […] The Bank needs to work with its development partners to reorient the Fast- Track Initiative to support improved learning outcomes, in parallel with the MDG emphasis on primary completion. (IEG-WB, 2006, p. xvii).

2006 also saw the publication of another deeply consequential research piece – the first that explicitly called for learning to become the centerpiece of the global development agenda. Under the unequivocal title, *A Millennium Learning Goal: Measuring Real Progress in Education*, the paper was published under the auspices of the Center for Global Development, and prepared by three World Bank-affiliated economists D. Filmer, A. Hasan and L. Pritchett. In the paper, the authors insisted upon the idea that the completion of basic schooling did not mechanically lead to the acquisition of universally necessary competencies. By drawing on PISA data from seven developing countries already meeting (or about to meet) universal completion of primary education, Filmer and colleagues concluded, "Many students complete their schooling well short of minimal competencies […] the majority of youth do not reach a plausible minimal competency level in mathematics, reading and science" (Filmer, Hasan and Pritchett, 2006, p. 9). On the basis of such findings, the authors of the paper derived a number of conclusions and policy implications that, over the years to come, would become greatly influential within development circles. Firstly, they considered it necessary to move beyond the education measures favored by the MDG framework, in order to focus on *actual acquired competencies;* and secondly, they argued that an outcomes-based approach to education quality might be a more appropriate catalyzer of change than an inputs-based approach.

In retrospect, the significance of such publications can hardly be overstated. Within the World Bank quarters, the need for a greater emphasis on learning and learning measurement continued to acquire institutional centrality, and the emphasis on learning was soon reflected in the programmatic work of the organization. One of the key episodes in this regard was the launch of the READ program, with the support of the Russian Federation[126]. The program ran from 2008 to 2015 and was oriented towards assisting

---

[126] The establishment of the READ program represents one of the first education projects undertaken by Russia in the area of international development after the fall of the Soviet Union (Takala and Piattoeva, 2017). Given its role as an "emerging" education donor, it remains difficult

low-income countries in improving their student learning outcomes through the establishment or refinement of student assessment systems. Beyond its impact at country level, the READ program was instrumental in consolidating learning assessment as an issue-domain central to the World Bank education mandate (World Bank, 2015). The process of institutionalization of the learning turn culminated with the launching of the World Bank Education Strategy in 2011, under the name *Learning for All*. In this programmatic document it was clearly stated that "Achieving Learning for All will be challenging, but it is the right agenda for the next decade. While countries can achieve rapid changes in enrollment rates from one school year to the next, it is much harder to make significant gains in learning outcomes" (World Bank, 2011, p. 9) and that "The new strategy focuses on learning for a simple reason: growth, development, and poverty reduction depend on the knowledge and skills that people acquire, not just the number of years that they sit in a classroom" (p. 25).

Importantly, the discursive shift initiated within the World Bank and adjacent academic circles also permeated other institutions. A case in point is the establishment, in 2007, of an FTI Task Team on the Quality of Learning Outcomes, as a result of a proposal put forward by the Russian Federation. The objective of the task team was to advise the FTI Steering Committee in relation to "guidelines and procedures in support of improved learning outcomes in partner countries" (Cambridge Education/Mokoro/Oxford Policy Management, 2010, p. 199). Importantly, as a result of the work of the task team, two additional indicators were added to the FTI Indicator Framework, thus becoming reporting requirements – namely, reading ability at the end of Grade 2 and at the end of primary education (Cambridge Education/Mokoro/Oxford Policy Management, 2010, p. 73).

Similarly, in the years to come, some of the major development funders began to shift their focus from access to learning outcomes, while embracing an outcomes-focused approach to education quality (cf. Shiffman, 2011, for a similar argument). This was particularly the case of the UK's Department for International Development (DfID) and USAID. The former published its education strategy for the 2010-2015 program under the banner, *Learning for All* (DfID, 2010) and has since made an explicit effort to generate political interest and commitment around the question of learning outcomes, both in the UK and internationally (see Berry, Barnett, and Hinton, 2015; and DfID, 2013). USAID, in turn, published an education strategy in 2011 (*Opportunity Through Learning*) that, again, identified the improvement of learning outcomes as an institutional priority. The report noted that "For USAID's education assistance to meet the challenges of the 21st century, our programs must be strategically aimed to achieve measurable and sustainable educational outcomes" (USAID, 2011, p. 6), and identified as its first goal, the improvement of reading skills for 100 million children in primary grades by 2015.

---

to discern the reasons that led Russia to focus on the area of learning outcomes and learning assessment – literature on Russian development assistance is notoriously scant.

Such developments did not go free from criticisms. A number of papers, published in the early 2010s, (sometimes devised as contributions to the post-2015 debate) were critical of the growing focus of learning outcomes, exhibited by a range of donors and education agencies. Barrett, for instance, drew attention to the fact that a sharper focus on learning outcomes was likely to translate into an increase in high stakes testing and a neglect of the procedural dimensions of education quality (Barrett, 2011a, 2011b). Similarly, Tikly (2015) argued that the dominant approaches to learning were informed by a reductionist understanding of learning that equated progress to an improvement of test scores. Schweisfurth (2015), in turn, also warned of the risks of conflating education quality and learning outcomes, neglecting pedagogical questions. The absence of an articulated debate on the notion of quality was equally highlighted by Alexander (2015), who also called for a greater focus on procedural variables. Remarkably, such contributions were critical of the limited, reductionist approach that had driven the bulk of education for development efforts over the last few decades, and drew attention to the problematic character of an excessive emphasis on access and completion ("quantity") at the expense of quality. In this sense, the diagnosis of the problem was similar to that advanced by the World Bank in the early 2000s. However, the solution proposed by these academic commentators differed significantly from that advanced by the learning turn theorists. Thus, the aforementioned papers called for a greater emphasis on all dimensions of education quality (including inputs and processes) rather than a simple shift in focus from access to learning outcomes.

Such discussion remained confined to the academic sphere. However, the significance of these contributions lies in the fact that they prefigured (and in many ways, informed) the emergence of a normative community alternative to that represented by the aforementioned proponents of the *learning turn*. A divide between those emphasizing quality and those primarily concerned with learning outcomes thus began to unfold. Even if the differences between the two approaches were originally a matter of nuance or emphasis, they ended up conforming two distinct communities of understanding, informed by different sets of ideas.

### 5.3.2. *Championing the access-plus-learning agenda in the context of the post-2015 debate*

While the question of learning outcomes had been gaining organizational prominence in the agendas of the World Bank, DfID and USAID for some time, the early stages of the so-called post-2015 process proved a crucial opportunity for the consolidation of this theme as a contentious area within the global education field. Put differently, while the learning turn is the product of a number of (predisposing) factors that precede the post-2015 debate, the growing prominence given to learning outcomes cannot be understood without reference to the deliberate and purposive action of a group of organizations in the lead-up to the adoption of SDG4. Such actors thus took advantage of the opportunity to revisit the priorities of the education-for-development field afforded by the post-2015

debate. Indeed, calls for a *learning goal* featured prominently within the first proposal for a post-2015 education agenda. However, it is interesting to note that such efforts were not spearheaded by the intellectual architects and early adopters of the learning agenda (i.e., the World Bank, DfID or USAID) but by a different set of organizations that, up until that point, had enjoyed limited visibility within the education-for-development realm, namely, Brookings and Save the Children – with crucial support from a range of foundations, among which the William and Flora Hewlett Foundation featured prominently.

One of the first and more influential advocacy initiatives which explicitly aimed at catalyzing a normative change within development circles was the Global Compact on Learning (GCL). Launched in 2011 and led by the Brookings Institution's Center for Universal Education, the GCE operated as a broad coalition bringing together a wide range of organizations (foundations, governments, experts, NGOs, etc.) seeking to "embrace, support and enact a policy agenda that focuses on access to quality and relevant learning opportunities for all children and youth" (CUE, 2011a, p. 2). The GCL was thus explicitly conceived as an instrument to galvanize attention and catalyze action (both within and beyond the education community) on the imperative to improve learning outcomes and to "refocus the global education discourse from solely getting children into school to ensuring that they learn while there" (CUE, 2011a, p. 1). Its first report thus called for a "paradigm shift" and made the case for an "expanded education agenda that centers the *goal of learning for all* as the new minimum threshold to which the education community must aspire" (CUE, 2011a, p. 3; emphasis mine).

Also importantly, the report framed the learning imperative with reference to equity challenges, drawing attention to the existence of huge disparities in educational attainment, and putting forward the notion of the "global learning crisis hitting the poorest, most marginalized children and youth particularly hard" (CUE, 2011b, p. 9). Such issue-framing would prove highly influential in the years to come, and represented a certain shift in the learning discourse. Until that point, much of the rhetoric surrounding the need for a learning focus had revolved around the expected returns (individual or collective), associated with improved learning outcomes – a rhetoric that essentially relied on the "logic of consequences", as per March and Olsen's (1998) classification of logics of action. The work of the GCL was an explicit effort to draw attention to the moral dimension of the learning turn – a narrative more clearly aligned with March and Olsen's "logic of appropriateness". While the equity narrative was not entirely new, the GCL was instrumental in placing equity *at the center* of the post-2015 advocacy efforts in favor of a learning turn.

The significance of the Compact also lies in the fact that the GCL began to devise the need for improved global learning metrics. The "visionary" nature of the GCL, in relation to the centrality of assessment, was recalled by a number of interviewees:

> Brookings had just published the report, A *Global Compact on Learning* in 2011. It was kind of an initial consensus among, primarily, donors and civil society groups in the global north and other think tanks on how we can start thinking on the learning crisis and the fact that kids are in school more today than ever, but not learning. And so one of the recommendations on that report was to have some better global metrics on learning. I There's a lot of different aspects in global education that are important. But we thought that if we could get a couple of good indicators on learning, that we could really kind of bring the field forward. (PRI4)

One of the main outcomes of the Global Compact was the establishment of Research Task Force on Learning, a group of experts chaired by Daniel A. Wagner (professor at the University of Pennsylvania and a highly regarded expert in the areas of learning and literacy). The task force, oriented towards developing a research agenda around learning and its embeddedness within international development efforts, was operative from December 2011 to December 2012. Despite its short life, the task force was instrumental not only in the definition of learning as an independent issue domain and an object of governance with well-defined contours, but also in creating a community of understanding around such themes, and in disseminating such ideas within academic circles and specialized circles (e.g., the task force has an open meeting in the context of the annual conference of the Comparative and International Education Society – cf. Wagner. 2014).

Additionally, the GCL gave way to a highly influential advocacy outlet – the Learning Metrics Task Force (LMTF). The LMTF was envisaged as a multi-stakeholder partnership co-convened by the Center for Universal Education (CUE) at the Brookings Institution and the UIS[127]. Much like the GCL, the LMTF was oriented towards fostering a "learning turn" in the education-for-development realm, taking advantage of the count-down to 2015. The project was explicit on the fact that:

> The overarching objective of the project is to catalyze a shift in the global conversation on education from a focus on access to access *plus* learning […] the task force works to ensure learning becomes a central component of the global development agenda and make recommendations for common learning goals to improve learning opportunities and outcomes for children and youth worldwide. (UIS/CUE, 2013a, p. i).

The task force kicked off in the early years of the post-2015 global education debate, and published its first report (UIS/CUE, 2013a) immediately before the Global Meeting of the Education Thematic Consultation – what presumably improved its agenda-setting potential. However, the real strength of the LMTF appears to lie in its inclusive character – the group brought together up to 30 organizations that met on different occasions from July 2012 to September 2013 (including regional and international organizations, donors, governments, bilateral organizations and NGOs). Additionally, the group benefitted from the contributions of nearly 200 experts and organized three consultative processes.

---

[127] The LMTF benefitted from the support of Dubai Cares, the William and Flora Hewlett Foundation, the John D. and Catherine T. MacArthur Foundation, the Douglas B. Marshall, Jr. Family Foundation and the MasterCard Foundation (UIS/CUE, 2013b).

Interestingly, according to one of the Brookings' officers engaged in the project, such willingness to secure a pluralistic and broad engagement was a response to the perception of the learning debate as potentially contentious:

> So yeah, so in the beginning it was Rebecca Winthrop [*CUE Director*] and a couple of others at Brookings, and we started talking with some of our funders and some people and basically people said, "This is a really contentious issue. It's going to be bloody". Somebody even said that word, "bloody". There's been these debates. No one's ever really been able to come to agreement. You have the teacher's unions on one hand, you have the World Bank, you have all these international agencies that don't necessarily play well together. Plus there's the millions of teachers and government workers and people that are really the ones who are going to make this happen who have been traditionally excluded from the conversation. So we were really pushed to… At Brookings, whenever we start a project, we start having conversations and consultations before we even develop our workplan to make sure that we get our finger on the pulse. (PRI4)

The partnership between the CUE and the UIS, in turn, was devised by Brookings as a strategic move to confer greater legitimacy to the project, thus rendering more acceptable its engagement with the post-2015 debate. In this sense, it is worth noting that the UIS appears to have been a second-best option. According to a UIS analyst, engaged in the project, Brookings had courted UNESCO before approaching the UIS. However, whereas UNESCO exhibited limited enthusiasm for the project[128], the UIS perceived the partnership as a highly strategic move, for it represented a way for the organization to contribute to the debate on the measurement of learning outcomes – with the financial and technical support of external partners[129]. It should, however, be noted that various interviewees have suggested that Brookings was always far more in control of the LMTF than the UIS:

> The reason they involved UIS and not UNESCO is because they were in contact with someone at head-quarters… So, it was part of an initiative initiated by Brookings, and they wanted to partner with UNESCO in order to get the legitimacy, because they know that UNESCO has the legitimacy to do something global. They contacted the director at head-quarters at the time, who was not the most approachable person, let's put it like that. And was not supportive of the project. So that's when they went to UIS. But initially it was intended to be a project with UNESCO, see? (UN7)

However, even if the engagement of certain constituencies was superficial, most interviewees concur that the LMTF was central in articulating and giving visibility to the access-plus-learning narrative – primarily by creating a sense of community and common purpose across different organizations. It is also important to bear in mind that the bulk of the LMTF activity revolved around measurement-related interventions, and in this

---

[128] Interviewees were unclear as to the reasons for this, an exception being the UIS analyst quoted above, who attributed UNESCO's limited receptivity to the temperament of a particular UNESCO manager at the time.

[129] As discussed in the following chapter, prior to this move the UIS had been forced to discontinue or reduce its work on the measurement of learning outcomes, against a background of severe budgetary constraints.

sense, the partnership was also instrumental in coupling the normative calls (i.e., the appeals in favor of a shift from schooling to learning) with a clear policy script – namely, learning assessments. In this sense, the role of the LTMF is central in explaining how the assessment program gained momentum and features on the agendas of a wide range of organizations in the development field.

According to a number of interviewees, while the legitimacy and credibility enjoyed by the LMTF owes much to its inclusive character and the presence of the UIS, its influence appears to be largely explained by the brokerage capacity, the network abilities and the charisma of one of its promoters – Rebecca Winthrop, the Director of Brookings' CUE. More specifically, Winthrop appears to have been particularly well-connected in different UN spaces. As a result, even if Brookings did not enjoy particular salience within the EFA strand of the post-2015 debates, it certainly had more sway in the UN workstream. Such dynamics are captured in the excerpts below:

> So I think this is an organization [*Brookings Institution*] that is listened to a lot in the UN. I think in many respects, what Rebecca [*Winthrop*] has done is fantastic, and she's a very, very influential person who is not part of the UN but is listened to quite actively. And the influence of Brookings, partly it has to do with Brookings itself. It's a very independent thing, right? It's based in Washington, DC, where a lot of these people hang out. But I am very impressed with Rebecca herself. A leader, she's been very, very good in terms of… she's not ideological. She's very fair, balanced, and she's very insightful. (PRI2)

> [*Interviewee probed about the growing centrality of learning in the post-2015 debate*] I would say the answer is very specific, and it can be traced to the Brookings Institution, and to Rebecca Winthrop, who was just very good at reading the room, and she was very attentive to the criticisms that were coming from Bill Gates, Bill Clinton, the big philanthropists, and she was very good at understanding that the education sector was being underfunded when compared to the health sectors […] So in a very skillful way, she said, "Well, let's start a dialogue, let's ask people around, let's have a conversation, let's make recommendations", and she collaborated with UNESCO people, with people from the Bank, but really, it was Brookings' initiative. (PRI6)

> What Rebecca did was, first of all, she approached Gordon Brown, and this is important, because at the end of the day, Gordon Brown is Gordon Brown. What can you say if a former Prime Minister says "This needs our attention?". And this is what Rebecca managed to do, these are her skills… Rebecca went all in, she used all her political mobilization capacity. (UN21)

Despite its centrality and visibility, the Brookings Institution was not alone in its efforts to push for a learning agenda in the context of the post-2015 debate. Another organization that was highly active in the intent to push a learning agenda was the NGO Save the Children. In a paper explicitly oriented at informing the post-2015 debate, Save the Children made a compelling case for an access-plus-learning agenda, while drawing attention to the global learning crisis and highlighting its equity implications – thus strengthening an already well-established narrative around the moral dimension of the learning agenda,

There are 130 million children in school who are not learning even the basics – a shocking figure masked by the focus in recent decades on getting more children into classrooms. As we look forward to the next set of global development goals, the focus needs to be on ensuring that no child is excluded – that every child, including the poorest and most disadvantaged, is both in school and learning […] In this paper, we argue that setting an ambitious global learning goal, as part of a post-2015 development framework, will be crucial to realising this vision […] Based on an assessment of the trends shaping the wider world and the changing educational context, Save the Children has proposed a post-2015 framework that tackles both the clear exclusion of children being out of school and also the hidden exclusion of children being in school, but receiving a poor quality education. Our proposals […] are underpinned by two core principles: *learning* and *equity*. (Save the Children, 2013, p. 4-5)

Save the Children thus engaged in a multi-pronged campaign strategy largely targeted at the UN strand of the debate. As reflected in the excerpt below, it relied on a combination of lobbying and advocacy efforts targeted at different audiences – including education specialists but also diplomats and delegations leading the OWG deliberations:

We've had a series of grants from the Hewlett foundation to push really for a strong learning goal and targets in the SDGs. So we started a grant I think in around 2013, to work on that, and we managed the grant that we had advocates in Ethiopia and Nigeria and also the UK working on that with us […] There were a lot of sign-on letters, and advocacy, policy statements that we used to share our strong learning goal. So there were several actions and statements that were made, in the course of that grant […] And every year before the UN General Assembly and during the different processes like the Open Working Group, we would put out policy statements. Certainly, before the [*inaudible - full?*] discussion of the Open Working Group on education, we put out a briefing. We have staff in New York, New York who are then lobbying the member states based on our briefing material and reports. (CSNGO16)

Importantly, and as reflected in the quote above, the engagement of Save the Children benefitted from the financial support of the Hewlett Foundation (see also Save the Children, 2015). This philanthropic organization was also one of the main funders of the LMTF initiative. Indeed, the question of learning outcomes and the access-plus-learning agenda had been elevated in terms of organizational priority within the Hewlett Foundation's portfolio since the mid-2000s[130]. While Hewlett's efforts had long focused on country-level interventions, in the run-up to 2015 the foundation initiated a grantmaking program explicitly targeting the global arena. As synthesized in the words of a foundation officer, such grants were oriented at supporting the incorporation of new organizations in the post-2015 education debate. As discussed below, the relevance of such development lies in the fact that, the "outsider" quality of some of the most

---

[130] The growing centrality of learning outcomes within the Hewlett Foundation's agenda can be traced back to the Quality Education in Developing Initiative, a grantmaking strategy established in 2008 in partnership with the Bill & Melinda Gates Foundation. This initiative was oriented at increasing awareness and public knowledge on the importance of learning outcomes, supporting instructional practices, improving learning levels on a broad scale and advocating a better allocation of education budgets so they contribute to maximize learning (Hewlett Foundation, 2008).

prominent advocates of the learning agenda are at the root of the animosity that learning-centered proposals generated within certain EFA circles:

> So one of the lines of work was to really take that global conversation to focus more on "What are the outcomes from education" and "How do we kind of improve those at a global level" […] We got involved, but from the perspective of looking at how we could support different NGO actors who are trying to bring people's voices into the process and similar society voices into the process, so, I mean, our main involvement was finding ways to support them to have a seat at the table in those conversations. So we supported a number of organizations like Save the Children… we wanted to also bring voices from the Global South to the conversation. So, I mean, I think our grantmaking was basically along those lines. We did also support through Brookings Center for Universal Education, so the Learning Metrics Task Force, again to kind of bring different experts and international organizations, national NGOs, governments together around a table to kind of just to build consensus around the importance of measuring learning and kind of what that means. (PRI11)

Despite the centrality of the Brookings Institutions and Save the Children in the advancement of the access-plus-learning agenda in the context of the post-2015 debate, it is important to bear in mind that the advocacy efforts of such organizations were ultimately consistent with the discursive shift experimented by the education-for-development community, described in the prior section[131]. In other words, the ideas disseminated by these organizations in the early stages of the post-2015 debate resonated with the broader environment and were echoed by more traditional development partners –most notably, the DfID and USAID. The prominence given to learning themes by these organizations in the post-2015 context was confirmed in interviews with representatives and spokespersons of such institutions, who listed learning as one of the priorities informing their contribution to the post-2015 debates (sometimes reflecting explicitly on the fact that this represented a departure from their prior agenda). The excerpts below, extracted from interviews with DfID and USAID informants, synthesize the priority given to the advancement of the learning agenda by these organizations:

> I think we have some clear priorities. For us, it was really important to get learning incorporated into the goal language and the indicator. And that was something different to what we stand for before. It was a priority to put an emphasis on learning. Yeah, it's a general priority for all to focus on learning, and then we wanted it to be reflected in the global goals. (DON3)

> So, USAID, overall, wants to have access for every child, and learning for every child. So we focus on access in countries of conflict or crisis. And then we focus on foundational

---

[131] The increasingly salience of the "learning goal" in the lead-up to 2015, can also be observed in the minutes and summary of the 2012 meeting of the International Working Group on Education (IWGE), an informal network of aid agencies and foundations oriented at developing a common understanding in support of education. The group is coordinated by the IIEP and regularly brings together representatives from UNESCO, UNICEF, the World Bank and major bilateral and foundations. The outcome document prepared as a synthesis of the 2012 meeting, which focused on the theme *From schooling to learning*, observed "No doubt there is consensus on learning as a goal", albeit it was noted that "No doubt there is consensus on learning as a goal" (Varghese, 2012, p. 48).

skills in the early grades. So we wanted to make sure... Access was never a really an issue, it was a carryover from the MDGs. We wanted to make sure there was something about learning, and if we could, we wanted to make sure that there was an indicator on the early grades. (DON4)

As noted in the following section, UNICEF was also identified by a number of CSOs and NGO representatives as a particularly vocal proponent of the learning turn. However, the centrality of learning themes within UNICEF's agenda remains something of a question mark. According to available documentation, the issue of learning outcomes was not particularly prominent in UNICEF's portfolio during the early stage of the post-2015 deliberations. UNICEF's position paper simply makes a cursory reference to the question of learning outcomes, noting that there is evidence of poor learning levels in developed and developing countries alike (UNICEF, 2013). Certain UNICEF's informants, however, suggest that the question of learning was indeed one of the organizational priorities in the lead-up to SDG4. A senior UNICEF official observed:

> When we were starting to talk at UNICEF about what's our UNICEF agenda for the SDG4 the sort of the push we were after was we do need to have something sharp on learning, which you probably know was quite controversial in some corners […] The equity agenda was already there but to turn an organization like UNICEF around to focus more on learning was actually quite a big [*inaudible - step?*] because it wasn't the norm to think about it in all programs in all places. (UN25)

A hypothetical explanation for such inconsistencies was offered by three different interviewees, who observed that, in the context of UNICEF, the learning agenda was embraced by individual bureaucrats (some of them holding high office) rather than at an organizational level. More specifically, the incorporation of a former DfID official in the UNICEF ranks, and the personal connection of this individual with one of the visible faces of the LMTF, were perceived to be at the root of UNICEF's veering towards the learning-outcomes and assessment agenda. While discerning the veracity of such assertions (or assessing their real significance) is beyond the scope of this dissertation, such observations are ultimately indicative of the general lack of clarity as to UNICEF's stance, in relation to the learning-centered approach. They are also indicative of the generalized perception that the learning agenda benefitted from the personal commitment of a number of well-connected individuals able to operate as idea brokers:

> I think if I was looking at what happened in UNICEF, I do think that [*deleted for anonymity purposes*] who came from DFID into UNICEF, really was a big part of the problem. I think she was pushing a very narrow agenda. And it comes with that sort of DFID perspective. (UN12)

> There are people with decision-making power… just to say that there are personal dynamics, and the strong relationship here of [*Brookings official – deleted for anonymity purposes*] and [*UNIFEF/DfID official – deleted for anonymity purposes*] … they took a liking to each other at the LMTF so when she was with DfID. And she really wanted learning and equity to be the center focus of the agenda. So personal dynamics they matter. (UN7)

It is, finally, important to bear in mind that efforts to promote a learning turn benefitted greatly from a focusing event[132] which originated within the EFA community and consequently, was much more difficult to ignore by the EFA pole of the post-2015 debate. More specifically, in 2012, the GMR published an estimation of the number of primary school-age children worldwide, who either had not reached grade 4, or had not reached a basic proficiency level, as established by cross-national learning assessments (EFA-GMR/UNESCO, 2012). The figure, estimated at roughly 250 million, had the virtue of synthesizing the growing volume of evidence on student achievement afforded by large-scale assessments. This was considered yet another proof of the existence of a "global learning crisis" – thus reinforcing the sense of urgency around the question of learning outcomes. The "250 million not learning" figure indeed became a powerful symbol, and diffused quickly to the point of evolving into something of a cliché – an oft-cited number that conveyed a powerful message that action on the need to act upon an unacceptable situation:

> I think the Global Monitoring Report which was released in 2013 if I remember correctly, it said that around 250 million kids could not read and write properly. So that was a very strong message, that even though you had included more children in schools, they did not learn very much when they were in schools. So I think that was a kind of a strong advocate for some more measurement to see if children actually were able to read and write and at least having the basic knowledge and basic skills. (CR2)

> I think, in terms of the different pieces of evidence that came to influence the debate the most, I would actually so that the 250 million children that were in school but not learning that the GMR put out I… that was really a figure that ended up being extremely influential, and I think that one was at the heart of the whole Save the Children stuff. It was quoted by many of the member states in the negotiations. (CSNGO5)

> So we have all of the international assessments, PISA, TIMSS and PIRLS. We have all the regional assessments, we have all of the EGRA assessments. So, a large body of research. And then you have major institutions in the field like the Global Monitoring Report coming out and saying 250 million children aren't learning. (DON4)

### 5.3.3. Learning and its discontents: "Dakar Wars" and the role of norm antipreneurs

As suggested above, the growing centrality of the learning-outcomes discourse within the development realm was never free of criticism. From the outset, efforts to promote a learning goal had been met with suspicion on the part of certain academic circles who drew attention to the risks of conflating quality and learning outcomes. However, and as the 2015 deadline approached, such concerns and misgivings transcended the academic community. The concerted effort to secure the salience of learning outcomes was thus

---

[132] Understood as an episode that brings less visible problems to the forefront, and allows for a re-interpretation of an already recognized issue, or gives additional prominence to an already existing problem, thus operating as a catalyzer of policy attention (see Kingdon, 1984).

met with little enthusiasm on the part of certain EFA partners – in particular, by representatives of the CSO and the NGO sectors.

The animosity exhibited by such constituencies appears to stem, to a significant extent, from the perception that the emphasis on learning outcomes would *inevitably* result in other themes being neglected – with learning outcomes eventually taking precedence over any other measure of education quality. Interviewees from such groups consistently expressed their fears over the possibility that the hype generated around the learning-outcomes theme would eventually have a narrowing effect on the education agenda. Such fears were, in turn, largely informed by the MDGs experience and the over-emphasis on access and completion themes that had characterized prior decades. Thus, a number of CSOs draw attention to the fact that the breadth of the education agenda needed to be vigorously defended against any attempt to privilege a specific theme at the expense of other issues. Such reasoning is captured in the quotes below, which exemplify the apprehension generated by the consolidation of learning outcomes as an issue area:

> So when we came to 2012 and 2013 and people started talking about what will happen in post-2015, we were concerned that the agenda could be reduced again, in the way that the MDGs had reduced ambitions of inclusive education agenda to just questions of access and gender parity. So we were very conscious at that time that the World Bank and UNICEF, particularly, with some backing from DfID, were keen to have a very narrow agenda, which would focus only on learning outcomes […] They assumed what we need is a single target goal around measuring learning, and the only easy measure is learning to read and write and maybe some numeracy. And so there was a big push for a very narrow agenda. Whereas our analysis, I would say, was that the problem was the narrowing of the agenda through the MDGs from what was an inclusive agenda […] And I think at some point people will look back and realize that it does more harm than good when you have a very, single-minded, narrow focus. (CSNGO3)

> If you think in terms of the Millennium Development Goals and if you think about education discourse 30 years ago or 20 years ago, everybody was talking about inputs, right? So the goals were in terms of enrollment and the idea was that if we get all children enrolled into school, if we have enough schools, if we have enough teachers, if we have enough materials then learning would happen. That link was [*inaudible*], right? Okay, now, everybody's talking about learning and learning assessments so they assume that making the data available will lead to change. That doesn't happen. We know from our experience, right? (LASS16)

At the same time, access was perceived to remain a pressing issue. Particularly in a context with increasingly high shares of displaced children and refugee population, different interviewees argued that access and completion could not be framed as mere "MDG carryovers". Other recurrent sources of concern were the testing and accountability implications associated with a greater focus on learning outcomes. Thus, different interviewees expressed fears over the possibility that the shift from schooling to learning would entail an intensification of standardized testing. Such a development was judged inappropriate given the potential implications in terms of teaching-to-the-test, curriculum standardization, the detrimental effect on children's wellbeing and the reinforcement of test-based accountability reforms:

186

There are several different issues that have been challenged, and which I think are difficult. One is a question about measurement. I think especially civil society, but also some member states are afraid that will narrow down to seeing education only as ... literacy or the ability to read and write and mathematics. But the broader agenda, this broader concept of uh, thus kind of a humanistic agenda, would lose out in a kind of a narrow, measurement agenda. So that has been one of the issues that has been there all the way I think. (CR2)

And I guess the pushback was coming particularly from strange quarters. So some of it was coming from civil society and obviously, there's a real worry about measurement being linked to accountability with teachers. And so the teachers' unions had quite a different [*inaudible*] as well because on the one hand they rightly say assessment is important for learning, but their fear with the SDGs was that it was going to drive more of a sort of a target culture around teachers, that teachers would be assessed by the grades. So there's a fear around how assessment will be used and probably some other things that they must feel very strong about as well. (UN25)

Additionally, some of the concerns and misgivings expressed by CSO constituencies (but also discussed in UNESCO circles) appear to stem from the perception that the learning agenda was external to the EFA community – the result of an outsider perspective. To some extent, concerns on the learning agenda were thus driven by the very identity and legacy of the organizations spearheading such efforts, or its intellectual architects. Since the World Bank was perceived as "removed" from a rights-based approach to education, the focus of the learning agenda was perceived by certain representatives as an attempt to bring an "instrumental" or economistic imprint to the global education agenda:

This focus on measurable learning outcomes, the read-write-count agenda, the way we see it, this was the agenda of the World Bank. This agenda, which was launched in 2010 and focused everything on learning outcomes, on standardized testing, this looked appalling to us. So our fight was to try to broad the agenda. To establish another paradigm. It was a struggle for the paradigm, because it was not an adjustment. What we wanted was not simply another target, or to adjust the target… No. We wanted a different paradigm. (CSNGO2)

People generally associate this working towards global metrics with the post-2015 agenda. But this was work that was happening even before the actual post-2015 consultations. But this was a reflection of that evolution of public education policy, which when from a focus, initially, from access only, to participation and completion, and then from that, the idea that completion in itself is not enough, that you actually have to learn. You know, like this simple formula of access plus learning, which developed back in 2010. So, you know, Brookings was pretty instrumental in that, and there was a lot of backing from kind of the Anglo-Saxon world, and the UNICEF world, and the World Bank, and you know, that kind of… the Washington-New York axis… I don't know what should we call it [*laughter*]. But you know, this was on the international rear screen anyway. (UN20)

So in between Dakar and this time, there were a lot of movements… DFID and USAID and the World Bank, all of them have come up with something called "Learning for all". It was about 2010. And they came up with this… some kind of publications. And all of them used a common phrase, "Learning for all". So kind of the moral or the trend is set by them. There is also the Learning Metrics Task Force who have been working on that… (UN13)

Interestingly, while the resistance generated by the learning agenda did not go unnoticed by the organizations actively supporting it, the reasons for such opposition were not always clearly taken into consideration. Criticisms were frequently dismissed as ideologically driven and born out of a *lack of understanding* or a *lack of realism.* Matters were compounded by the fact that the organizations more actively campaigning against the "learning agenda", were perceived as peripheral and consequently, portrayed as a vocal minority. The absence of organizational ties or recurrent interactions between the main proponents and the detractors of the learning agenda prevented the emergence of a dialogue required to find common ground:

> The learning factor is contentious. And some countries and some organizations are concerned about having more indicators on learning and what that means for… sort of narrowing what we mean by learning in school and also sort of test-writing children as either achieving or not achieving. And so, yeah, there were clearly sort of generally heartfelt concerns about whether or how the right way to capture learning in the global goal. And we were conscious of these concerns. I think there were lots of countries, there were lots of organizations that shared our desire to get learning indicators in there. But there are also some that felt strongly that it shouldn't be. (DON3)

> I mean I do sort of find that the Global Campaign for Education's agendas sometimes are less grounded in the reality of a particular country. And it's not hard to understand that maybe the advocacy point on testing in Chile, for example, it should be "Maybe do less of it, because it's hurting our children"; but in a country like Uganda it should be "Do more of it because we don't know if they're learning". And so that contextualization I think is so important. (UN25)

To be sure, there were attempts on the part of certain NGOs and CSOs to engage with the LMTF initiative in an effort to prevent its thinking from taking a reductionist turn (or from leading to a "testing craze"). One CSO organizer defined the strategy as "having someone on the inside to whom one could send messages" (CSNGO10). This was indeed consistent with the participatory, pluralistic *ethos* to which the Brookings Institution aspired. However, the success of this rapprochement strategy appears to be mixed. On the one hand, the early work of the LMTF (and particularly the first report, UIS/CUE, 2013a) was largely considered a meeting point, for it privileged a broad understanding of learning while explicitly avoiding its equation to literacy, numeracy or the so-called foundational skills. However, and as the work of the LMTF progressed, certain CSO representatives felt their engagement was being instrumentalized, in order to confer greater legitimacy to the initiative, whilst their input was having limited impact – for it did not succeed in preventing the LMTF from privileging a "narrowed" understanding of learning:

> So myself and [*deleted for anonymity purposes*] decided that we would join the process for a period of time, make sure that we fought within that process, and to challenge efforts to say that all you need was a very narrow indicator around literacy and numeracy. And actually, the LMTF initial reports are much, much more balanced because as soon as you look at the evidence it's very clear. Of course, you have to look at the breadth of learning, of course you need to be factoring in… and you have to recognize that measuring is not

the answer to everything, and if you only value the things that you can easily measure you have unintended consequences. There was some effort to recognize those issues of complexity. But, in the end, I think after that two years of involvement I realized that they were using my involvement as part of a sign of the credibility and inclusiveness of the task force. And so I distanced myself from it. And I wrote a piece which basically challenged this weird obsession. (CSNGO3)

So at least we were able to bring people together and say, "Are there certain things we can agree on and move forward?". But some members felt that there was still some things that they felt passionately about that didn't get addressed by LMTF, and went back to the old way of doing things, which was, like, "Let's publicly write an article about all the things". And so what we tried to do was get people at least around the table pushing the train in the same direction. And I think we got at a lot of people, even if we didn't get everybody and some people fundamentally still had issues. (PRI4)

In this sense, the LMTF debates prefigured a conflict that would only escalate as the post-2015 debates progressed. By most accounts, the normative conflict between those pushing for a learning-centered education goal and those promoting a broader, quality agenda crystallized in the context of the Education Thematic Consultation led by UNESCO and UNICEF, and particularly in the global EFA meetings held in Dakar in March 2013 (see section 5.2.2 above). The meeting was shaped by an ideological cleavage that had been long in the making, thus revealing the existence of opposing views as to the priority given to learning outcomes vis-à-vis other education themes or proxies of education quality. The sense of animosity that permeated the whole meeting led a certain CSO to dub the meeting "the Dakar Wars":

A very important meeting that took place in Dakar in 2013, we called it the Dakar Wars, it's an analogy with Star Wars. Because was happened in the meeting was terrible. But it was a victory, it was a success what we managed to get from the meeting. (CSNG12)

GCE kind of got its act together and it progressively got its act tougher… the Dakar meeting was a wakeup call, [*we understood that*] unless we get our act together, we will get this learning outcomes testing agenda, therefore we need to… we need to actually start acting and influencing. (CSNGO10)

Thus, a number of developments led a range of CSOs to perceive the meeting as an attempt to impose a narrow agenda – a new version of the MDGs in which the focus on learning outcomes, was to replace the focus on UPE and gender parity:

UNICEF and the Bank were working on a document for the targets, something very closed to the read-write-count agenda, and they called this paper "the non-paper". And in the Dakar meeting, there was a televised message from Gordon Brown, and he proposed a super-goal, "Learning for All". So what was the expected result from this meeting? It was to reach a consensus that "Learning for All" should be the education goal. That was the plan! And during the meeting, it was Brookings who took the floor, all the time. All the time. I couldn't believe the lack of transparency. Because Rebecca [*Winthrop*] was speaking all the time. Whereas the GCE had like one minute and a half. And Member States they had a roundtable! So that's important, the fact that Member States were not playing a role… It was more an individually-driven thing, see, with the High-Level Panel of Eminent People, the Global Compact… (CSNGO2)

In light of these developments, a number of CSO organizations (and most notably, Education International, the International Council for Adult Education and the Global Campaign for Education) adopted a watchdog role – an active strategy focused on ensuring that the learning agenda was kept at bay. More specifically, representatives of these organizations report having made significant efforts to guarantee that the new agenda would not focus exclusively on the question of learning outcomes (and would not be framed as an access-plus-learning agenda):

> There was a point where we met and we said… let's boycott. Because they were, "Ok, we have a learning for all goal, so what do you think should be the targets?" And we were saying "No, who said this was the goal? We are going to do this, we are going to do this exercise, because if we do the exercise, it means that we are legitimizing the goal, but we haven't reached a consensus". And we insisted, "There's no consensus". So the struggle was to make a point of saying, "You are trying to impose a false consensus, this is not true". And a colleague approached me and told me, "I heard a rumor, I'm hearing that they are actually pushing this in the final plenary". And then, in the final plenary, they were saying, "So we have a consensus that the super goal is learning for all". And the Senegalese government was saying we only have 5 minutes for questions because we have to close the session. So the civil society raised, we requested to speak, and we said "Firstly, there is no consensus around this, secondly, how can we have an education goal that does even have the word education in it?" So that was a success. (CSNGO2)

> There was a small pre-meeting in Bangkok the week before the Dakar big meeting […] I guess the tensions started coming up around the SDGs, around the Education 2030 Agenda, started really coming to the surface at that point. For me, the strongest memory I have is the struggle over quality versus learning. And I came back feeling pretty proud of myself. My first time of blocking the UN statement until corrections are made. Because they were going after a learning agenda. And we were making sure education with a slightly more holistic definition that is not just about standardized learning outcomes. I kind of held up the statement for two hours. The poor person who was trying to compile it was so upset with me [*laughter*] (CSNGO10)

Therefore, over the months to come, learning-centered proposals made limited progress. Such developments cannot be disconnected from the evolving architecture of the post-2015 education debate. They appear to owe much to the greater prominence acquired by UNESCO and the EFA Steering Committee during late 2013 and early 2014, as well as to the greater formalization of the debate:

> We broadened the process. The earlier process, again, was very opaque. It was following the MDG Track, with a few agencies and donors and experts sitting in New York crafting all of this and spinning a whole thing around it. And I think civil society pushed a broader process, to hear out a wider set of voices, bringing new perspectives, especially from the poorer countries, bringing genuine researchers, bringing the voice of young people, of women, etc. And after all, in the EFA process also and the whole EFA dynamic, the discourse around education as a human right was very strong already. So when we open that process, when we pushed so that UNESCO made sure that that process became real and became an active player in the whole dynamic, then that eventually [*inaudible - held sway?*]. (CSNGO12)

However, such efforts continued in the context of the OWG deliberations, particularly in light of the efforts made by Save the Children to promote a learning-centered goal, as the excerpt below suggests. It is, however, relevant to note that, according to Wulff's[133] inside account of the OWG debates, the "battle between education and learning" that pitted EI [*Education International*], the GCE and the ICAE [*International Council for Adult Education*] against Save the Children, was often "difficult to understand" (2020, p. 46). Such observations are ultimately indicative of the difficulty to transfer education-specific debates to the broader development or UN community.

> And there you had EI, the Global Campaign, the Adult Education Council, as the primary organizations that really didn't want to frame anything really in terms of learning outcomes but wanted to keep it focused on quality and a broader notion of quality which - it goes back to the type of inputs that you need to have in an education system to have any quality, but it also goes back to our fear of numeracy and literacy becoming the only thing that will then be somehow focused on, and what that then in extension means for standardized testing and narrow curricula and so on and so on. And there I would say Save the Children was our kid of primary opponent, in the sense that they did a lot of lobbying for a learning goal. And Save the Children is also kind of organized enough to be quite a difficult organization to work against, they were very present in all of the negotiations and everything- Plan International was more Save's side than ours, but they were not as present throughout the process […] But I think what we did manage to do was often shut Save the Children up in some of the joint statements that we did, because they ended up being the only or one of few organizations that were demanding certain language on that. And then we could kind of over-rule them just with reference to a majority of the organizations signing up to the statement being against it. (CSNGO5)

Ultimately, it can be argued that such CSO constituencies acted as *norm antipreneurs* in the sense advanced by Bloomfield (2016) – that is, as actors engaged in *preventing* and *resisting* the consolidation of a given normative change. More specifically, within the norm antipreneurs spectrum, the CSOs opposing the rise of learning outcomes as the linchpin of the SDG agenda can be characterized as *creative resisters* (Campbell-Verduyn, 2016) – in that they did not defend the normative *status quo*, but accepted that some degree of change was necessary. Thus, such actors did oppose those normative entrepreneurs willing to frame learning outcomes as the centerpiece of the post-2015 agenda, but conceded that some change was necessary – as the focus on schooling and the emphasis of the last decades, had proven problematic. Thus, as a compromise solution, a number of CSOs advocated for a broader quality framework – one that combined attention to learning outcomes with explicit attention to inputs and processes. Additionally, and in order to prevent a reductionist understanding of learning (i.e., one equating learning outcomes to numeracy and literacy), they pushed for the inclusion of the *relevant* adjective behind the first reference to learning outcomes:

> I think that quality as well ended up being a convenient word that everyone could somehow agree on – you could kind of hide the tensions around learning outcomes under the kind of broader umbrella of quality education. (CSNGO1)

---

[133] The author of the piece, A. Wulff, was one of the negotiators representing Education International in the context of the post-2015 debate.

I mean, the whole learning outcomes issue was an issue that never was resolved. What we ended up doing was often add "relevance" in front of "learning outcomes" and that was something we did together with the indigenous peoples major group, because that was something that we felt would at least broaden the scope a little, even if of course it's one of these fluffy concept that has to be kind of defined to have any meaning. (CSNGO5)

Ah, the quality bit! In the Framework for Action, we wanted to make sure that the definition for quality was beyond measurable learning outcomes. And we succeeded. But we could not eliminate the reference to learning outcomes. So it is here, but it is re-dimensioned. (CSNGO2)

As the aforementioned discussion suggests, the negotiation of SDG4 proved a turning point in the life cycle of the learning norm – a crucial test that led to its refinement. While the final wording of SDG4 draws attention to the question of learning outcomes, this is not portrayed as an overriding priority. The adoption of SDG4 thus represents a shift from a schooling-centered agenda to an expanded education agenda, rather than to a learning-centered agenda. In this sense, the negotiation of the SDGs entailed a normative shift, albeit not the shift originally expected by the advocates of the learning agenda.

The relative influence of the aforementioned norm antipreneurs in such refinements cannot be seen as straightforward. Certainly, the wording of SDG4 represents a clear and explicit endorsement of a broad agenda and a holistic view of education quality. Both the final goal and the interim proposals (including the outcomes of the Education Thematic Consultation, but also the TST issue brief and the bulk of the OWG proposals) are framed as education agendas, rather than as learning agendas (cf. Tables A.1 and A.3 in Annex 2). While learning outcomes are portrayed as an area in need of greater attention, such observations are "balanced" with references to other indicators of education quality, relative to inputs and processes. As noted by Sayed and Moriarty (2020), the SDG4 is indicative of a quality turn, rather than of a learning turn (see Unterhalter, 2019, for a similar argument on the *broad* understanding of quality that eventually prevailed in the formulation of the SDG4 targets).

However, such developments cannot be exclusively explained as the result of the counter-advocacy labor undertaken by CSOs or certain EFA constituencies. Rather, the changes might respond to the expansive and aspirational nature of the SDGs. Ultimately, the general willingness to avoid a limited agenda (similar to that epitomized by the MDGs) rendered virtually impossible the advancement of a simplified, narrow "learning goal" such as that promoted by the architects (and early adopters) of the learning-outcomes agenda. Put differently, a counterfactual analysis suggests that, even in the absence of a number of norm antipreneurs contesting the learning agenda, it is unlikely that such an agenda would have made significant progress – as the framing and the overall narrative surrounding the SDGs was not conducive to the identification of a narrow set of priorities. In this sense, the broad and holistic approach to education quality eventually embraced by the SDG4 is not necessarily (or not exclusively) the result of the counter-advocacy campaign described above, but of a broader transformation of the practice of global goal-

setting. This is reflected in the shift from a simple and succinct agenda (the MDGs) to a consensual and expansive development framework (the SDGs).

### 5.3.4. The learning agenda beyond SDG4

As discussed in the previous section, the negotiation of SDG4 was settled in favor of a broad, holistic understanding of education quality. However, the consensus reached in the lead-up to the adoption of SDG4 proved to be a fragile one. Soon after the adoption of SDG4 in late 2015, a number of proposals oriented at reinstating learning outcomes at the center of the global education agenda were put forward by different organizations.

The Learning Target (LT)[134] launched by the World Bank in October 2019 features as one of the most influential and explicit efforts to reinstate learning as the centerpiece of the education agenda. Thus, less than five years after the adoption of the SDGs, the World Bank introduced a new target, namely, "By 2030, reduce by at least half the share of 10-year-olds who cannot read". The LT was conceived as an operational target expected to guide the World Bank's operational work in low- and middle-income countries, however, it soon acquired a universal quality, for it was promoted and announced as a "top development priority", along with the twin goals of ending extreme poverty and advancing shared prosperity (World Bank, 2019a, p. 20)[135].

In this sense, the rapport of the learning target and SDG4 was described in rather ambiguous terms. On the one hand, the report introducing the LT noted that the new target was aligned with the spirit of the SDGs – in that both the LT and SDG4 recognized the need for improved learning outcomes, and ultimately aspired to widespread, high-quality education. On the other hand, the report suggested that the SDGs might be overly aspirational in nature, excessively difficult to reach and, consequently, unlikely to generate the necessary momentum to effectively guide political and policy action. This line of reasoning is captured in the excerpt below:

> The Sustainable Development Goals—and specifically Goal 4—show that the international community now recognizes these problems. The new focus on learning embodied in the SDGs is a significant advance over the Millennium Development Goals, which promoted in- creases in access but not improved learning. *But nearly five years into the SDG era, it is time to take stock and make course corrections.* Given the depth of the learning crisis, there are reasons to question whether the current targets under SDG 4 are feasible, or whether new intermediate targets are needed to spur concrete, focused action. (p. 9, emphasis mine)

---

[134] Sometimes referred to as the Learning Poverty Target (LPT).

[135] The World Bank's continued emphasis on the need to reinstate the improvement of learning outcomes as a development priority can also be observed in a range of recent publications, but particularly the *World Development Report 2018, Learning to Realize Education's Promise* (World Bank, 2017). While adressing a variety of themes, this publication insisted on the need to prioritize learning and highlighted the risks posed by the learning crisis.

The LT thus appears to reinterpret SDG4 by prioritizing a subset of objectives and nudging countries towards focusing their efforts on a specific goal, deemed more feasible. In this sense, the LT entails a certain comeback of the spirit of the MDGs, i.e., the preference for a narrow, simple goal – judged more likely to equip the global development community with a clear sense of direction. As a consequence, and even if the LT was framed as an effort to *complement* SDG4, the initiative soon raised suspicion within the global education community and was perceived in certain quarters as an instance of counter-targeting. In particular, those organizations more heavily involved in the negotiation of SDG4, frequently voiced criticisms on the duplicative and reductionist effect of the LT initiative, while noting that the new target lacked the legitimacy enjoyed by SDG4 on account of the inclusive and consultive character of the post-2015 process. Additionally, the launching of the LT was perceived to be an indicator of the World Bank's limited ownership of the SDGs – a development reminiscent of the dynamics that had characterized the tail-end of EFA and considered to be at the root of the limited traction enjoyed by the Dakar goals. The following quotes from UNESCO senior officials and a senior OECD observer of the process synthesize such a line of reasoning, and are indicative of the cross-cutting character of such perceptions:

> They [*the World Bank*] are a co-convener agency for the EFA and for the SDG Steering Committee… but I have to say, they were not very prominent. At all, really. They took part in the meetings, but… And I think this has always been the case! I mean, all through EFA it was the same, and from Jomtien 1990, the World Bank has never really engaged in the international development targets […] I think they like to have the flexibility of, you know, setting their own targets, setting their own indicators… And I think several of the senior education managers find working in the UNESCO-led forum, they find it deeply frustrating. And I think they feel that they don't need it. I have a very strong impression that they feel that they don't need this. The relationship with their clients is so strong that the global framework, doesn't matter. (LASS3)

> I think it's always been the issue from the beginning, what would be the World Bank's role with the SDGs and the nature of the interactions. And the most recent developments, moving to one learning target or whatever it's called, it's not unexpected. I would say, because in a way, again, I think it's part of the whole process of narrowing of the outcomes of learning to learning being defined as a performance on a standardized test of literacy, essentially that. (UN30)

> The World Bank is basically kind of disengaged from this global Steering Committee around the new agenda. And money is not everything, but it's a significant part of the whole equation. So, if the World Bank is having their separate discussion with countries, with absolutely no consultation with other partners except the country, and it's got to do with money. (UN20)

Remarkably, the World Bank is not the only organization whose programmatic priorities have recently switched focus towards a learning-centered agenda. The strategies of both GPE and UNICEF place learning at the centerpiece of their respective education agendas. Thus, under the unequivocal rubric *Every Child Learns*, UNICEF's Education Strategy 2019-2030 is organized around the notion of "the right to learn", and the report explicitly states that "the Strategy outlines the shift towards a greater focus on improving learning

outcomes" (UNICEF, 2019, p. 4). In coherence with this vision, the report identifies three main goals which, once again, revolve around the notion of learning, namely, "(1) equitable access to learning opportunities; (2) improved learning and skills for all and (3) improved learning in emergencies and fragile contexts" (UNICEF, 2019, p. 24). A similar shift in focus can be observed in GPE's Strategic Plan 2016-2020, entitled, *Improving learning and equity through stronger education systems.* In this case, "Improved and more equitable learning outcomes" also feature as one of the three main goals identified for the years to come, again in response to the global learning crisis (GPE, 2018). The process through which learning (and specifically, the improvement of learning outcomes) have eventually become an organizational focus of these IOs, is beyond the scope of this dissertation. In this sense, the impact of the SDG4 negotiation process on this shift in emphasis, remains an open question. Additionally, the ultimate impact of such strategies on the organizational priorities of these institutions, remains to be seen. However, such developments are indicative of the prominence of the improvement of learning outcomes on the global education agenda. This suggests that, despite the ambiguous success made by the learning turn in the context of the post-2015 debate, a normative shift, aligned to the learning turn is well underway.

Another development that is indicative of the ultimate success of the learning turn can be observed in the support for the introduction of a *lead* or *flagship indicator* on the part of different organizations focusing on the progress made in relation to learning outcomes. One of the most prominent supporters of this idea is DfID. Through the *Better Education Statistics and global Action to improve learning* (BESTA) program, DfID has allocated £7.8 million to "Establish new global learning indicators, including a new flagship learning indicator for education" (DfID, 2018, p. 1). As part of such efforts, DfID has been supporting the UIS in the development of a composite indicator, combining access and learning data – the so-called Children Not Learning (CNL) indicator. This measurement is expected to operate as an advocacy tool – a simple figure conveying a strong and clear message, and consequently more likely to raise awareness, galvanize political action or attract additional funding for the education sector (UIS, 2017b).

Interestingly, calls for the development of a global lead indicator appear to have their origins in the work of the International Commission on Financing Global Education Opportunities (commonly known as the Education Commission) – a platform bringing together experts in a variety of fields and national representatives, chaired by the UN Special Envoy for Global Education and former UK Prime Minister Gordon Brown, and oriented at "invigorat[*ing*] the case for investment in education, to bring about a reversal in the current underfunding" (UNESCO News (2015), unnumbered)[136]. Thus, one of the key recommendations of the Education Commission as a means of encouraging action

---

[136] The Commission was created in the context of the Oslo Summit on Education for Development, held in July 2015 and convened by Prime Minister Erna Solberg of Norway, President Michelle Bachelet of Chile, President Joko Widodo of Indonesia, President Peter Mutharika of Malawi and UNESCO Director-General Irina Bokova.

and investment in the education area, as well as placing greater emphasis on learning outcomes, was the creation of a lead global learning indicator (Education Commission, 2016). As captured in the words of one of the Education Commission leaders below, the lead indicator is explicitly devised as an advocacy tool, oriented at shifting policy and public focus into learning:

> The Education Commission came out with some very high-level recommendations but also recommended the number of strategic initiatives, which were what they were. And one of the strategic initiatives was to move forward global lead indicator learning. This would be kind of built into or building on the SDG process, the SDG indicators but would serve as sort of the advocacy number kind of… our school children have been the lead indicator for education for the past 10 years about but wanted to transition to something that's more focused on learning and that would serve to really call people's attention to the crisis in education and galvanize support as well. (PRI10)

It is, however, interesting to note that, very much like the World Bank's LT, such proposals have been met with a certain skepticism on the part of CSO constituencies, as well as on the part of those organizations, more openly committed to preserving the centrality of SDG4, as the only legitimate global agenda. Once again, such misgivings do not stem exclusively from the very content of the Education Commission's proposals, but also from the process through which such recommendations have been crafted, and from the embeddedness of the Commission within the broader SDG4 architecture:

> There are question marks. And I think we'll see some continuing battles on this side to the Education Finance Commission. This was the sort of Gordon Brown and the Norwegians and others which tried to reintroduce the idea of a single lead indicator that has tried to create a parallel architecture with a high-level UN special representative who would report directly to the UN Security Council and Human Rights Council. And this is really an attempt to create a job for Gordon Brown. We are very concerned that that is an effort to create parallel to what has been agreed as their follow-up process with the Education 2030 steering group. (CSNGO3)

**Wrapping up**

This chapter has examined the process through which a new global education agenda was crafted, from two different but complementary perspectives – one relative to the very architecture of the negotiation, the other relative to the content of the new agenda. In relation to the former, the chapter has found that the coexistence of two workstreams (one led by the UN, the other by EFA) often proved a source of friction, as a result of the disparities between the bureaucratic culture of the UN sphere and the deliberation routines that characterize the EFA architecture – a divergence that, in turn, reflects different understandings of what qualifies as a democratic and inclusive debate. In relation to the debate on the content of the new agenda, the chapter has found that, while global goal-setting exercises have been documented to operate as opportunities for the consolidation of normative shifts, such processes are not necessarily as linear or straightforward as some accounts might suggest. Indeed, attempts to institutionalize the learning outcomes norm

as the centerpiece of the post-2015 debate did not go unchallenged. Norm-entrepreneurship efforts were actively resisted by certain strategically-located constituencies, which forced a certain tweaking of the norm in order to render it politically palatable and secure consensus. Overall, the negotiation patterns identified in the chapter suggest that, while the formulation of the SDGs was a deliberative and participatory process, these qualities should not be equated to an absence of conflict or hierarchy. Thus, even when the negotiations relied on horizontal decision-making structures, different actors were unequally equipped to take full advantage of them. Also relevantly, consensus-building exercises frequently relied on language ambiguities, rather than entailing a genuine process of ideational convergence or rapprochement.

# No dataset left behind?
# Reporting on SDG 4.1 and the politics of data integration

## Introduction

The previous chapter dealt with the negotiation of SDG4 and its associated target. It also addressed the quantification efforts that paralleled this process – that is, the *conceptualization* of SDG4 indicators, which culminated in the adoption of the global and thematic reporting frameworks. The present chapter resumes the process of indicator-making by addressing the very next step – the *production* of indicators, understood as a collective process involving not only a few theorizers but also a variety of data producers, collectors and curators (cf. Chapter 3 for the theoretical underpinnings of this distinction). In keeping with the dissertation's emphasis on the learning-outcomes debate, this chapter focuses on the production of the global indicator corresponding to Target 4.1[137], namely, "Indicator 4.1.1 Proportion of children and young people (a) in grades 2/3, (b) at the end of primary, and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex" (UNGA, 2017) [138].

This is in fact perceived by a range of education stakeholders as the SDG4 "leading" indicator, and is arguably the one that represents the biggest challenge for the education community. This is because its production requires the development of new instruments, methodologies and standards that have historically defied consensus. Thus, despite the broad agreement on the centrality of learning measurement, the production of Indicator 4.1.1 has been far from straightforward and conflict-free. At the time of writing, and five years after the adoption of the SDGs and the Education 2030 Framework for Action, the reporting protocol still remains under discussion[139]. Difficulties encountered in the process have not been exclusively technical in nature, but also stem from the difficulty of

---

[137] That is, "By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes" (UNGA, 2017).

[138] Note that, while in the Global Indicator Framework this is the only indicator for Target 4.1 in the Thematic Indicator Framework this indicator is accompanied by six additional indicators. See Chapter 5 for a discussion on the interlinkages between the global and thematic frameworks.

[139] Even if an interim reporting strategy has been defined, and there is growing consensus regarding data collection and harmonization standards, such agreements retain an unstable and fluctuating quality, and its real strength will only be seen over the long haul. However, the objective of this chapter is not to analyze the robustness or adequacy of the indicator production efforts – let alone to speculate how the process will play out in the future. Rather, it aims to gain insight into the mechanics of global quantification, and to understand the transformative impact of an inherently collaborative project on the agendas and relationships of the partaking organizations. In this sense, the absence of a final agreement is not judged to be an impediment – conversely, it helps to set the focus on the structural dynamics shaping the process rather than on an end-product whose specifics might simply reflect contingent circumstances.

reconciling the multiple (and sometimes conflicting) institutional interests, priorities and assumptions of the different IOs and data-producers partaking in such efforts. If the negotiation of the SDG4 learning targets was a story of consensus building (or at least, conflict suppression), the quantification of such targets is one of fragmentation – a process in which many of the strategic ambiguities exploited in the post-2015 process have become readily apparent.

This chapter aims to gain insight into the mechanics of global quantification and, more specifically, to identify and make sense of the organizational dynamics behind the production of Indicator 4.1.1. The chapter delves into the process through which indicators are coupled to specific data suppliers, statistical routines and reporting standards. Attention is paid both to high politics and technical debates, since both are equally relevant to making sense of the different steps involved in the assembly of global statistics. The chapter is structured as follows. The first section provides a historical overview of the organizational arrangements set up by the UIS in an effort to produce global learning data. The second section discusses the tension between in-built comparability and country ownership as two basic principles shaping SDG4 reporting efforts, and provides an overview of the early stages of Global Alliance to Monitor Learning (GAML) (2015-2017) – a period in which such a trade-off became particularly manifest. The chapter follows with the analysis of the later stage of GAML (2017-2020), with particular attention to efforts made by the UIS to bargain with a range of stakeholders and to craft a consensus in a fragmented field. The last section of the chapter analyzes the evolution of the UIS' position within the realm of learning metrics, with particular attention to the internal and external conditions that have made this transformation possible.

## 6.1. Setting up an institutional infrastructure

In 2015, with the adoption of the Education 2030 Framework for Action, the UIS was ratified as the official source of cross-nationally comparable data on education (i.e., nationally-reported, globally-aggregated data). Later, the Institute[140] was designated the custodian agency[141] for most SDG4 global indicators. The transition into the SDGs era therefore represents a critical moment for the UIS in that it enjoys unprecedented levels of visibility. The production of Indicator 4.1.1 in particular can be characterized as a turning point in the history of the Institute, as it placed the UIS in a central position in the

---

[140] Note that the terms *UIS* and *the Institute* are used interchangeably in this chapter.

[141] The notion of custodianship refers to responsibility on "(i) methodology development for new indicators; (ii) supporting the increased adoption and compliance with internationally agreed standards; (iii) collecting data from countries; (iv) compiling internationally comparable data-series; and (v) strengthening national statistical capacity (IAEG-SDG 2017)" (UIS, 2018, p. 11). In practice, this means that UIS is requested to provide the UN Statistical Commission with data on the selected indicators, on an annual basis (IAEG-SDG, 2017).

production of globally comparable learning data – an area in which the UIS had long struggled to affirm its leadership and expert authority.

It is important to bear in mind that production of globally-comparable learning data exhibits a number of particularities that set it apart from other indicators produced by the UIS. Such specificities are relative to the nature and number of data sources involved in such efforts. Thus, the production of global databases on learning data relies necessarily on the data produced by organizations other than statistical offices and national governments[142]. A great share of learning data is produced by Cross-National Assessments (CNAs) whose rapport with the UIS is much more tenuous and less institutionalized than the relationship between the Institute and national agencies. The engagement with learning assessments has long been a challenge for the UIS, which has primarily relied on administrative data for the production of its core indicator set[143].

Such dynamics explain why, historically, the UIS has never approached the production of global learning metrics as a *solo* endeavor. Rather, it has sought the collaboration of leading assessment organizations and experts as a means to secure both the validity and the legitimacy of global metrics. This first section provides an overview of the different collaborative spaces set up by the UIS in its quest for global learning data. Attention is given to the early attempts (even if, as discussed below, they met with little success) as well as to the most recent efforts made in the context of SDG4 – specifically, the establishment of GAML.

*6.1.1. Historical precedents*

As documented by Addey (2018), UNESCO (and, by extension, the UIS) can be considered a late-comer to the learning assessment phenomenon. In 2003, and in consideration of the then recently approved EFA indicators, an expert meeting has held in UNESCO "to develop a working definition of literacy and a conceptual framework for literacy assessment" (p. 399) – a move that, as noted by the author, cannot be dissociated by the fact that UNESCO's statistical leadership was severely questioned, not the least given the OECD's growing prominence in the realm of assessment (through the

---

[142] More specifically , for the elaboration of global datasets, the UIS has historically relied on the data collected by national statistical offices and education, science and culture ministers and national governments, who are expected to report their administrative data according to the standards established by UIS standards. However, it should be noted that the UIS is not directly in contact with OECD countries. Data from such countries are collected through the so-called UOE procedure – according to which Eurostat collects data from EU Member States and the OECD does the same for the rest of OECD states (Cussó, 2019). In this sense, the UIS's "global outreach" might be more nominal than real.

[143] A similar challenge has been posed by the incorporation of data produced in the context of household surveys led by other IOs (e.g., UNICEF's MICS) or development partners (e.g. , USAID's Demographic and Health Survey) (UNESCO, 2018).

implementation of PISA but also of adult literacy assessments[144]). One of UIS' first inroads into the area of learning assessment was the creation of the *Literacy Assessment and Monitoring Programme* (LAMP), initiated in 2003 and oriented at measuring adult literacy. However, the project suffered by a number of political, technical and institutional setbacks, and "never gained sufficient global prestige" (Addey, 2018, p. 400).

Assessment efforts were redoubled in 2008, when the UIS started to stimulate some degree of cooperation (or, at least, communication) among international and regional assessments. In collaboration with UNESCO'S Education Sector, the UIS undertook a series of efforts to "promote a systematic UNESCO-wide approach to the assessment of learning outcomes" – a strategy that came to be known as *Assessment of Learning Outcomes* (ALO) (UIS, 2008, p. 4). The ultimate objective of this initiative was to produce internationally comparable statistics on learning, and the main strategy to achieve this included the development of an international bank of cognitive items which was to be used in existing testing initiatives. Thus, with the establishment of ALO, the UIS moved beyond its initial attempts at *producing* assessment data and focused on its role as *data harmonizer.*

One of the main outcomes of ALO was the establishment of an *International working group for assessing and improving quality learning* – a deliberative body gathering the main global and regional organizations involved in learning assessments. The initiative thus brought together representatives and experts from the International Association for the Evaluation of Educational Achievement (IEA), the OECD's Programme for International Student Assessment (PISA), the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), the *Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN* (PASEC), the *Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación* (LLECE), as well as experts from the World Bank, the EU and the FTI (UIS, 2009).

This was seen as major step, largely as a consequence of the animosity between the IEA and the OECD. A former UIS senior official recalled with amusement that, in the meeting that inaugurated such efforts, someone wanted to take a picture of a panel session bringing together some senior OECD and IEA figures – for they had not been seen together in a long time (UN21). Reaching a compromise between these parties soon proved an improbable outcome, and the group never succeeded in generating comparable information. According to a former UIS manager (UN24), international assessments had little incentive to collaborate, were afraid of "losing control", and invoked a number of technicalities preventing them from engaging in data and item sharing.

---

[144] Cf. Grek (2009), for an overview of the process through which the OECD succeeded it consolidating itself as the "the pre-eminent global organisation for developing and analysing comparative international educational performance data" (p. 28) largely as a consequence of the success of PISA. See also Grek (2013) for an overview of the OECD's efforts in the international adult literacy studies and the transcendence of such developments – in that they represented the first instance of international testing with an explicitly comparative dimension.

Later, the UIS engaged in a new effort to produce cross-nationally comparable information on learning outcomes, and to coordinate what was becoming an increasingly competitive and crowded space (UIS, 2011). The new initiative, known as the *Observatory of Learning Outcomes* (OLO), was launched in 2011. However, financial constraints[145] prevented the UIS from conducting substantial work on this area (UNESCO, 2012), and the project was suspended in 2013 as a cost-efficiency measure and in order to "preserve resources for core services" (UNESCO, 2013a, p. 19). Also during 2012, the UIS was in conversation with the leading assessment organizations (most notably, IEA, LLECE, PASEC and SACMEQ), and even reached an agreement with them to start comparing reading skills at the end of primary education. The initiative was discontinued in 2013 "due to decisions by partners to focus exclusively on their assessments" (p. 73).

The limited success found by the UIS in the monitoring of learning outcomes, as well as the budgetary restrictions faced by the Institute during this period, are two of the key factors behind the Institute's eagerness to join forces with the Brookings Institution through the creation of the Learning Metrics Task Force. The initiative was seen as a privileged opportunity for the UIS to contribute to the debate on the measurement of learning outcomes that was then emerging in the context of the post-2015 development agenda. In addition, and as noted by a former UIS manager, it was perceived as a way to depoliticize an issue perceived as potentially contentious (UN24).

In 2014, and in the light of the growing prominence of the "quality measurement agenda", the UIS resumed its efforts to produce comparable data on learning outcomes. To this end, the *Learning Outcomes Advisory Group* was created. Devised as an expert committee, the board drew together representatives from the leading organizations in the assessment realm (OECD, IEA, SACMEQ, etc.), but also specialists from IOs (World Bank, GPE) and, more importantly, a number of independent specialists with recognized expertise in the assessment area (i.e., researchers not affiliated with any organization and without any vested interests – or only tenuously so). Thus, and in contrast with ALO, the OLO group was devised as a more expert-driven forum (UIS, 2014).

However, despite efforts to create a neutral space not driven by organizational interests, the harmonization of assessment data was again deemed an ill-advised project. The OLO recommended instead a reinforcement of the UIS clearinghouse function. In response to this, the UIS engaged in the production of a Catalogue of Learning Assessments – a

---

[145] By the early 2010s, UNESCO was experiencing severe financial constraints, largely as a consequence of the membership dues withheld by the United States upon the admission of Palestine as a UNESCO Member State in 2011 (cf. Hüfner, 2017). In the case of the UIS, this translated into a 24% reduction in total general income over an 18-month period – a loss resulting mainly from the reduced allocation from UNESCOs Regular Program, but compounded by a reduction in the contribution of the Government of Canada. In is in this context that the UIS redoubled its fundraising efforts, signing agreements with new donors and looking for additional partners (UNESCO, 2013a).

project initiated previously but discontinued in practice given the lack of resources (cf. UNESCO, 2014f). According to a former UIS analyst, the recommendation was motivated by the perception that data produced in the context of different assessments were inherently difficult to compare and defied harmonization.

> So for some reason, they couldn't agree. So you put these experts in a room, for days, a few times a year, and they could not agree what is it that they could and could not collect. After endless discussion it was decided in the end that it would not be possible to report learning outcome levels in a database. So those were the recommendations of the Learning Outcomes Advisory Board [...] The advisory board was [*saying*], "You know, you can't compare the learning outcomes levels of all these different assessments, but what you can do is to compare the characteristics of the assessments, the population, how are the results disseminated, how are they analyzed, whether they use criterion-reference or norm-reference, are they low-stakes or high-stakes". All those characteristics about assessment that would emphasize that the results cannot be compared, that the data could not be compared. (UN7)

The UIS thus put considerable effort in the elaboration of a Catalogue of Learning Assessments which mapped existing national and international assessments, along with public examinations, and was expected to fulfill several functions. Marketed as "the only open data sources of internationally comparable information about learning assessments" and framed explicitly as a public good, it simultaneously aimed to support policy-makers in making decisions about assessment, and to serve the donor community by assessing the quality and use of the programs they were funding (UIS, 2015).

### 6.1.2. The Learning Metrics Partnership

By mid-2014 UIS' efforts in the area of learning outcomes were focused on the creation of the Catalogue of Learning Assessments. However, given the direction taken by the post-2015 debate and the recommendations made by the TAG (cf. Chapter 5), the UIS started to anticipate the need to produce some form of globally comparable learning data. The international monitoring of learning progress was increasingly perceived as a goal that could no longer be postponed. Thus, in April 2014, the UIS and the World Bank convened a meeting in Montreal with the objective of defining criteria to monitor reading in primary education. It is in this context that the possibility of developing a baseline or common metric by linking regional and national assessments started to take shape (UNESCO, 2014f).

This objective eventually materialized in the creation the *Learning Metrics Partnership* (LMP)[146]. The LMP was conceived as a joint initiative of the UIS and the Australian

---

[146] The project is considered by many to have originated in the context of the Learning Metrics Task Force, to the point that a number of interviewees thought of it as a sort of prolongation or spin-off of the LMTF – although the UIS put some effort in separating the two projects. The lack of a clear demarcation between the two initiatives is captured in the words of an NGO officer, who, in reference to the Montreal meeting discussed above, noted, "The meeting was early in

Council for Educational Research Centre for Global Education Monitoring (ACER-GEM), supported by the Australian Department of Foreign Affairs and Trade (DFAT). The partnership aimed to "develop a set of nationally and internationally-comparable learning metrics in mathematics and reading, and to facilitate and support their use for monitoring purposes in partnership with interested countries" (UIS/ACER, 2014, p. 1). More explicitly, the initiative aimed to establish a metric[147] against which existing assessment could be plotted, and to support countries in using their assessments to report on global targets based on this tool. In this sense, the activity and products envisaged by the LMP were very much an embryonic version of some of the work that would subsequently be developed in the context of GAML.

However, the LMP initiative was short-lived; a change in the leadership of the UIS in 2015 precipitated an early termination of the project[148]. According to different interviewees, the incoming UIS Director (Silvia Montoya, replacing Hendrik van der Pol) decided to put an end to the partnership with ACER in order to reaffirm the centrality of the UIS and its visibility in the creation of globally comparable learning data. As noted by a UNESCO official:

> So DFAT was going to give a contribution of, I think, 5 or 7 million to the UIS. And I think that there was some kind of condition attached to that, which was that ACER would play the main role within this topic... So Silvia was new, had just came in, and was like "What?! What's going on? This is the topic that interests me, I want to give leadership to it, I don't want this to be sold to someone else immediately!" (UN12)

At the same time, the move was also prompted (or at least, facilitated) by the reservations expressed by some regional assessments, who perceived the initiative as excessively top-down in nature (for they had hardly been consulted despite the centrality of their input in the construction of a global scale), and were reluctant to share their items with ACER. A former UIS analyst noted thus that the LMP had fallen apart after a particularly contentious meeting in which regional assessments perceived that they had been unduly sidelined:

> And then we had a meeting and the first two days went very well, and the third day ACER was invited to present their thing, and that's where the trust was lost in the UIS by the regional assessments. In that, it seemed that it was a *fait accompli*. That it was already a

2014… it might not have been an official LMTF meeting. So I went to Montreal for these measurement meetings and at one of them at least, we were told very specifically, 'This is not an LMTF meeting' and we were like 'What? This is a conversation with the same people with the same topic?' So I think [*there was*] a little confusion about which was called what and when" (CSNGO6).

[147] Understood as a *common scale* capturing different stages of educational progression in a given dimension – not a *common assessment*. As noted by different ACER and UIS informants, the confusion of both concepts is at the root of a number of misunderstandings and suspicion of the initiative.

[148] Formally, subsumed under GAML according to UIS (2016a).

deal, an agreement. That UIS would work with ACER and would just compare the countries regardless. The little hiccup was that for ACER to do their job, the regional assessments had to submit a bunch of information about the items, and the difficulty levels, psychometrics... and of course, that didn't work. (UN7)

Finally, it should be noted that the limited support found by the LMP within the learning measurement community was also a consequence of the leading role played by ACER in this venture. Some interviewees were under the impression that the LMP was being used by ACER as a "shop front" to market its services, and that ACER was instrumentalizing global targets to create a niche of its own in the assessment market – especially after its central role as a PISA contractor came to an end after the 2012 survey. Remarkably, some ACER representatives were quite explicit in their willingness to carve a niche for themselves within the development realm. An ACER analyst noted that the advent of the SDGs and Education 2030 was a "very fortuitous set of circumstances" in that "it became a place in which our thinking and our work was able to develop and expand and become more sophisticated". He noted similarly that over several years, ACER had been putting considerable effort into expanding into the development realm in order to compensate for their loss of centrality within the PISA program:

> [*After 2012*] So we've had a lot of expertise that was available and we've been channeling quite a bit of that in support of the international as its developed its thinking around the Sustainable Goals for Education since then […] Over the last three or four years, we've put on more senior staff to lead and guide developments of relationships with those international development-oriented agencies, so the World Bank, DFAD, the Australian aid program, the various United Nation agencies such as UNESCO and UNICEF, and so on. We've developed strong relationships with a wide range of players in the international community, largely building our own experience and expertise developing an increasingly strong reputation because of the quality of the work that we've been able to do with those agents. (LASS10)

Since the centrality of ACER within the LMP did not help in bolstering the credibility of the common-scale proposal as a sufficiently disinterested one, it ultimately provided the UIS with a further reason to discontinue the LMP in favor of a more inclusive and neutral space.

### 6.1.3. The Global Alliance to Monitor Learning

By the end of 2015, a provisional thematic indicator framework had been approved and the Global Indicator Framework was underway (see Chapter 5 for an overview of such efforts). Given the unprecedented emphasis put on education quality and learning outcomes, the production of global learning data became an organizational priority (and an inescapable responsibility) for the UIS.

It is up against the urgent need to report on SDG4 targets, and the failure of the LPM venture, that the Global Alliance to Monitor Learning (GAML)[149] was created under the auspices of the UIS. Announced in late 2015, it was devised as a group bringing together a wide range of stakeholders (most notably assessment agencies, education-related IOs and INGOs, and national authorities) with the objective of reaching an agreement on the specific measures and tools necessary to monitor learning in the context of the SDGs (Montoya, 2015). The group was formalized in 2016 and celebrated its first reunion with the support of the World Bank, who hosted the meeting. Since then, GAML has met on seven occasions. For clarity purposes, and since this chapter makes multiple references to such encounters, Table 6.1 below offers a timeline of GAML open meetings.

**Table 6.1. Meetings of the Global Alliance to Monitor Learning**

| Meeting | Date | Venue |
|---|---|---|
| *1st meeting* | 11 May 2016 | Washington, DC (US) |
| *2nd meeting* | 17-18 October 2016 | Washington, DC (US) |
| *3rd meeting* | 11-12 May 2017 | Mexico City (Mexico) |
| *4th meeting* | 28-29 November 2017 | Madrid (Spain) |
| *5th meeting* | 17-18 October 2018 | Hamburg (Germany) |
| *6th meeting* | 27-28 August 2019 | Yerevan (Armenia) |
| *7th meeting /GAML webinar* | 10 June 2020 | On-line |

Source: Author's elaboration.

The creation of the GAML was largely welcomed by the education community on account of its inclusive character as well as an explicit recognition that the challenge at hand was simultaneously technical and political (Montoya, 2015). Thus, the change that resulted in the combined effect of a new UIS Director and the creation of a new space, was portrayed by a number of interviewees as a positive development – an opportunity to start with a clean slate. This was especially true of those informants less well-versed with the decades-old fight between the IEA and the OECD, and who were very critical of the paralyzing effect of such disputes (which they perceived to be increasingly irrelevant and essentially driven by personal antagonism). Additionally, the advent of GAML was perceived as creating more space for data producers other than large assessment consortia. As an NGO officer put it:

> It's sort of a breath of fresh air because the rest of us have been sitting at these tables for years now, literally! […] We've been sitting here for four years going around and around and around, with vested interests and new models, and not moving forward. And she [*Silvia Montoya*] has been like "Okay, this is what we're going to try next and this is what I need you to do". And I think it has been great, from my perspective. Maybe others are frustrated, but from where I'm sit, I'm clear how I can help her. (CSNGO6)

GAML originally aimed to "address the need to expand and improve the production of measures of learning outcomes and skills to effectively monitor the targets established in

---

[149] Initially announced as *Global Alliance for Learning* (GAL) (see Montoya, 2015).

the Sustainable Development Goal 4" (UIS, 2016b, unnumbered). However, the objectives of GAML were never restricted to SDG4 monitoring – that is, the development of tools and strategies to measure progress towards SDG4 learning-related targets. As discussed in the *Concept note* published in 2016, and as emphasized by the UIS Director in one of her first GAML presentations, GAML objectives were in fact quite broad and included (*a*) the production of technical guidelines and good practice in large-scale assessments (relative to survey design, test administration, sampling, and quality assurance procedures) oriented towards supporting international, regional and national data producers; (*b*) strengthening the learning assessment capacity of Member States through capacity-building, knowledge-sharing practices, and the support of a resource-mobilization strategy aimed to increase funding for large-scale assessments (UIS, 2016c; Montoya, 2016).

In this sense, it is relevant to note that GAML was not exclusively concerned with the enhancement of the international comparability of learning data, but also with an improvement in the quality and availably of learning data, as well as with the sustainability of assessment efforts. Thus, the work of GAML not only aimed to serve the international community, but also to support the work of Member States through the promotion of an effective use of data and an evidence-based approach to policy-making (Montoya, 2016). As discussed below, this multiplicity of mandates and anticipated data users is central in explaining many of the trade-offs faced by GAML and the UIS over the following years.

Largely as a consequence of its ad hoc nature, the composition and structure of GAML, and even the role of the group within the Education 2030 architecture, remains something of an elusive subject of study. The Alliance[150] has indeed been characterized by an evolving structure and a changing composition.

Thus, in line with its self-professed "open and participatory" nature, GAML membership has, since the start, been open to any individual or organization interested in contributing to its work. However, the first GAML meetings were mainly attended by international agencies, development partners, research organizations and foundations with a global scope (UIS, 2016d). A wide range of individuals with different institutional affiliations were critical of the absence of country representatives at GAML meetings. Some of them even casted doubt on the real interest of many countries in the production of globally comparable data, noting that the needs of the international community and those of Member States were not necessarily aligned. An NGO officer noted that:

> You know, it's a really teeny-tiny group of people that used internationally comparable data, and they happen to be all in the same room, and what's missing in the room it's national offices. There no national assistants, there are no national ministry of education. Not one voice. And there are various people around the table raiding it, raising it again, myself included! (CSNGO7)

---

[150] Note that the terms *GAML* and *Alliance* are used interchangeably in this chapter.

The sentiment is also clearly articulated in the words of a representative of a non-profit organization, who argued that "it sounds like the countries don't quite understand what GAML is doing and like they don't understand the technical solutions that are being proposed or the big ideas circulated at GAML […] there's no evidence that I've seen that countries would embrace the proposal that so far have been floated at GAML". Since 2018, in response to such concerns, an explicit effort has been made to incorporate country representatives and these representatives have been acquiring a greater centrality in GAML meetings.

> The first meeting of GAML only had technical partners around the table. And I would say to an extent understandable. It's such a technical question, you can't quite open the discussion to everybody, you need to find a common language, and share common objectives, and then you can present that to the political level. Now, this seems to have open up a little bit more... I guess it's a realization that things can move quickly, so you don't want to present a solution completely all of a sudden. (UN17)

Some interviewees noted that this move was not only necessary to secure the UIS' authority *within* the education for development community, but also *externally* – since the involvement of countries conferred the Institute with a much-needed legitimacy in the eyes of the IAEG and the UN Statistical Commission[151]. One of the analysts in charge of a regional assessment noted:

> The role of the countries is also important because the US has to defend its own work against the Statistical Commission, the Inter-Agency group, and so on. So it's important to have the support of the countries… to make sure they are convinced this is the appropriate strategy. Also it's kind of risky to advance in a unliteral way, you don't want to present this to countries. (LASS13)

Relatedly, the governance structure and decision-making protocols were progressively defined as a means for the UIS to demonstrate a commitment to transparency and inclusion. Such changes need to be understood as a response to the criticism faced by GAML during the early days of the initiative. The reliance on self-selection dynamics, the lack of clarity regarding the governance of the Alliance, and the limited formalization of decision-making procedures soon came under the critical scrutiny of a number of members. The latter was particularly problematic for those attendees representing large international bureaucracies, and who were consequently less willing or capable of making snap decisions, as captured in the excerpt below:

> Part of the reason people were anxious in those meetings was that decisions were just kind of being made in the moment. Papers hadn't necessarily been presented in advance, people didn't know what they were deciding on. There were no clear recommendations.

---

[151] As discussed in the previous chapter, the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs) is an ad hoc forum established by the UN Statistical Commission (UNSC) in 2015. The group was tasked with the development of the indicator list for the Global Reporting Framework, which was adopted by the UN General Assembly in July 2017. The work of the group has continued its work as a facilitator of the implementation of the framework.

> This is not a point of blame. It was just a situation. And you saw some organizations, particularly in the UN, extraordinarily anxious that they didn't feel that they could make decisions in the moment, like they have their processes, they need to get back to their organizations […] Because it felt like, you have Silvia [*Montoya*], who's very comfortable because she's so smart and so quick in kind of brainstorming in the moment and they trying to get a solution. But you have other organizations who kind of freak out and say, "no, wait a minute, I can't just make this decision. I have to take this ack to my headquarters" (DON6)

An issue alluded to by a number of interviewees was the sense of confusion created by the existence of multiple debates (some of them preceding the GAML-led process), which only a few savvy players were able to navigate and fully decode – an NGO representative compared these dynamics to "the telephone game"[152] (CSNGO6). Another officer from a non-profit development organization noted:

> It's hard to tell what's going on. I mean, I'm not actually privy to all conversations, so I can't speak from inside knowledge, but it seems is more of a process where there's something that's being presented, and then, you know, there's an audience that sits there and consume the information that's being presented, but many people are not sure yet how they would voice their reservations or what the process would be to build consensus. (PRI3)

Largely in response to such concerns, and in the context of the GAML 2nd meeting, an agreement was reached to set up a "light structure" organized along a UIS-based Secretariat, a series of Bureaus and Co-chairs, and a number of task forces (UIS, 2016e). Such structure was later refined in a *Governance and organization note* issued in 2017 and in preparation for the 3rd meeting. According to this note, the technical work would fall on a thematic task force, but the inputs and recommendation prepared by them should be discussed and endorsed by plenary meetings. On the basis of such decisions and discussion, the Secretariat (served by UIS) is expected to "finalize recommendations and report to the TCG[153]" (UIS, 2017c, p. 10). The note thus made it explicit that GAML was expected to "decide and elevate technical solutions to the TCG about all learning outcome-related indicators" (UIS, 2017c, p. 18). Finally, one of the key innovations put forward by this note was the creation of a Strategic Planning Committee (SPC) oriented towards fostering greater levels of transparency and overseeing the role of the work of the UIS Secretariat. The SPC was composed of a UIS-appointed Chair, the UIS Director, and a number of UIS-proposed Co-chairs representing different constituencies. The SPC Terms of Reference, however, were not adopted until GAML's 4th meeting in late 2017

---

[152] A children's game in which players, sitting in a circle or line, have a message whispered in their ear, and they then pass it along to the next person. Typically, the final message ends up differing significantly from the original message, as each one of the players needs to reconstruct and reinterpret it – and errors accumulate.

[153] Technical Cooperation Group – a platform convened by both UIS and the UNESCO Education Sector's Division of Education 2030 Support and Coordination. The TCG has a political mandate to develop and debate SDG4 thematic indicators – cf. Chapter 5.

(UIS, 2017d). Since then, GAML's structure and workflow has remained relatively stable.

Overall, an effort has been made to secure a more democratic environment – and despite some reservations, the move has proved successful in reinforcing both GAML's and UIS's authority. Taking stock of such changes, a researcher involved with GAML's efforts noted "There's been a broad consultative process... it's almost a peer review process. And they had originally started down, I think, a quite slightly different path... and they moved over again something much more open" (PRI5). However, it is important to bear in mind that, as will be discussed below, the democratization of GAML remains an unfinished process – by mid-2019, some participants were still unclear on the real *locus* of decision-making, and reported being under the impression that the actual distribution of power differed slightly from the one laid down in the official documentation.

## 6.2. The (relative) value of in-built comparability

Efforts to produce SDG 4.1 data have been shaped by the existence of competing visions of what constitutes good assessment data. While in-built comparability and technical sophistication are properties of paramount importance to some actors, others fear that an excessive emphasis on such features might end up preventing the Large-Scale Learning Assessment (LSLA) program[154] from realizing its policy-planning potential, by incentivizing countries and donors to privilege a specific subset of assessments administered externally and unlikely to realize the full potential of the learning metrics.

This section provides an overview of the early stages of GAML – a period in which such tensions became manifest. The section starts with a brief presentation of the dilemmas and trade-offs posed by SDG4 monitoring needs – particularly those relative to the multiplicity and heterogeneity of data sources. It follows with a description of the SDG 4.1 reporting strategies initially considered by the UIS, as well a description of the hybrid strategy that was eventually privileged. The rest of the section is devoted to analyzing the key episodes behind the preference for this hybrid approach – the process through which some courses of action were discarded whereas other ideas were preserved or at least repurposed. This section therefore documents the demise of in-built comparability as the

---

[154] As signaled in the previous chapter, one of the ideational repercussions of the learning turn has been the coupling of the *learning-crisis problem* with the *learning-measurement solution*. The idea has found major resonance in different IOs, who have incorporated the improvement of learning assessment systems as one of their programmatic priorities. Thus, Large-Scale Learning Assessments (LSLAs) appear to have acquired the status of a consensual policy script within the development realm, largely as a result of their potential as a policy tool (that is, a tool useful to spur government action but also to inform domestic policy-making – for instance by informing regulatory and resource-allocation activity).

overriding priority governing the production of global learning data – and the rise of country ownership as an essential principle.

*6.2.1. One indicator, multiple data sources*

Even if the creation of GAML was largely perceived as a positive development, by 2016 the production of globally comparable learning data remained as challenging as ever. This was the result of a number of factors – starting with the fact that the notion of learning proficiency defies simple definition or desires for a neat categorization. Thus, because the notion of learning has no obvious demarcation and there are no self-evident ways to measure it, the production of learning metrics remains an inherently contested endeavor. The debate has proven all the more contentious in the context of the SDGs, for there is growing awareness that the quantification efforts associated with the global goals are highly consequential since they have the potential to elevate certain definitions to universal categories.

Concerns regarding the concept-measurement gap were in turn compounded by difficulties relative to the availability and quality of data sources. This second group of challenges included the limited coverage of existing datasets (i.e., the absence of data on learning outcomes for a number of countries), the limited comparability of existing data, and the existence of multiple and overlapping datasets that required prioritization.

It is important to bear in mind that, by 2016, there were a variety of CNAs, but no consolidated methodology to equate and harmonize this data. This reflected the fact that, as captured by Table 6.2, such assessments are notoriously heterogeneous in terms of targeted domain and grades, and in their design, purpose, mode of assessment, sampling, methods for score estimation, etc. (for an overview, see Treviño and Órdenes, 2017).

**Table 6.2. Overview of cross-national assessments**

| Target potentially informed by the assessment | Grade or target population | Name of the assessment | Domain – Reading (L) or Numeracy (N) |
|---|---|---|---|
| -- | 1 | EGMA | N |
| | 1 | EGRA | L |
| 4.1.1a | 2 | PASEC | L, N |
| | 2 | EGMA | N |
| | 2 | EGRA | L |
| | 3 | LLECE | L, N |
| | 3 | EGMA | N |
| | 3 | EGRA | L |
| | 4 | PIRLS/ePIRLS | L |
| | 4 | TIMSS | N |

| 4.1.1b | | LaNA | L, N |
|---|---|---|---|
| | | PILNA | L, N |
| | 5 | SEA-PLM | L, N |
| | | LLECE | L, N |
| | | PASEC | L, N |
| | 6 | SACMEQ | L, N |
| | | PILNA | L, N |
| | | LaNA | L, N |
| 4.1.1c | 8 | TIMSS | N |
| | 15 y.o. | PISA | L, N |
| | 14-16 y.o. | PISA-D | L, N |
| -- | -- | ASER, Uwezo (5-16 year-olds) | L, N |

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

*Notes:*
- Grey: regional assessments; Light Grey: international assessments; White: assessments of foundational skills and population-based assessments.
- Includes only those domains relevant for global reporting purposes.

The possibility of integrating data sources other than CNAs constituted another major debate. By 2016, a growing number of countries had a National Assessment (NA) in place, but it remained unclear how such data could be used for comparison purposes. It is important to bear in mind that, whilst NAs are not particularly amenable to comparative purposes (at least in a raw form), the use of NAs for global reporting purposes enabled the rapid expansion of the coverage of Indicator 4.1.1. In addition, discarding NAs as a valid source of comparable data was perceived as a risky move since it runs against country-ownership principles. Also, it might create a perverse incentive for countries to invest *exclusively* in external assessments which are unlikely to realize the full potential of learning metrics, especially when it comes to education planning and policy design (for they lack the frequency and granularity necessary for such purposes). It should be noted here that, as Addey and Sellar have highlighted, countries already have a number of strong incentives for participating in CNAs[155] – their participation being driven by a wide range of political, economic, technical and socio-cultural rationales. In addition, low- and middle-income countries in particular are frequently encouraged (or requested) by some donors to take part in CNAs (Addey and Sellar, 2017, 2019).

Of course, implementing an NA and joining a CNA are not mutually exclusive options – and in fact, CNAs are sometimes portrayed as a first step or a model for NAs (for instance,

---

[155] Note that the authors refer to *International Large-Scale Assessments* (ILSAs) – a denomination extensively used by the specialized literature. For the purposes of this dissertation, I use the term cross-national assessments in order to signal the commonalities and differences with national assessments.

PISA has been promoted by the OECD as an opportunity to improve technical capacity at the national scale – cf. Addey and Sellar, 2017). However, in a resource-constrained context where assessment practices are nonexistent or remain at a nascent stage (as is the case of many low- and middle-income countries), national policy-makers, development partners and donor circles have to make a decision as to which assessment modality should be prioritized.

*6.2.2. Towards a feasible reporting strategy*

Tacking stock of such trade-offs, the UIS prepared a first list of the different options available to produce comparable data for Target 4.1, identifying the political, technical and financial implications of each of them (UIS, 2016a). The range of options was thus compiled in a report published in March 2016, just before the 1st GAML meeting. In the report, the UIS signaled its preference for one particular subset of options, which were identified as the most pragmatic and cost-effective (UIS, 2016f). However, over the following years, the remaining options kept resurfacing in the context of GAML debates, albeit under different names. The undying nature of such proposals owes much to the organizational interests associated with them – that is, the fact that certain data producers saw in them an opportunity to reassert their own centrality in the global efforts to report on SDG 4.1. Table 6.3 below summarizes the different options considered by the UIS to produce global reporting metrics, as well as the political technical and financial costs associated with each of them according to UIS' estimations.

**Table 6.3. Options for the production of global reporting metrics – Target 4.1.1**

| Option | Description | Costs and challenges |
|---|---|---|
| *Establishment of a new test* | Development of a new assessment specific to a target population and learning domain, to be implemented in all countries | High economic costs
Long implementation process
Political challenges associated to countries' limited buy-in |
| *Backward-linking cross-national assessments* | Linking results from existing assessments on the basis of a common item pool mapped against an agreed content framework. | Political challenges associated to the (anticipated) resistance of existing assessments to share items and student responses. |
| *Forward-linking cross-national assessments* | Redesigning future cycles of CNAs in order to render them comparable – primarily through the establishment of a common framework (i.e. new items on anchorage packages to be incorporated in future assessments). | Political challenges associated to difficulties in reaching a consensus and fostering collaboration among existing CNAs |
| *Forward-linking national assessments* | Extension of the *Forward-linking cross-national assessments* option by linking national assessments to the common framework. | Political difficulties associated to countries' limited willingness to engage
High technical cost associated to harmonization efforts |

| | Adaptation and expansion of a specific CNA so that for each grade, all countries participate in same assessment. Need to establish a "regional module" to avoid floor effects and include contextually-relevant items. | |
|---|---|---|
| *Enhancing cross-national assessments or "Adopt a test"* | | Political difficulties associated to countries' limited willingness to engage |

Source: Author's elaboration on the basis of Montoya (2015), Treviño and Órdenes (2017) and UIS (2016a, 2016f).

As noted above, the UIS initially identified two of these options as the more feasible. More specifically, the forward-linking of CNAs and the enhancement of CNAs were ranked as the most viable and promising options (UIS, 2016a, 2016f). However, over subsequent years, both options proved to be highly challenging from a political point of view. Thus, during the period 2016-2019, the UIS tinkered with a number of reporting strategies with the objective of devising an option acceptable to all parties involved. The profusion of reporting protocols, mapping exercises, prospective studies and concept notes produced, is testimony to the multi-pronged strategy pursued by the UIS over this period. Thus, rather than favoring a particular course of action, the UIS explored a variety of reporting and harmonization strategies and commissioned work to a wide range of parties – hence securing a role for all data producers without privileging any of them.

The interim reporting strategy put forward by the UIS[156] is thus informed by a hybrid approach that combines multiple data sources. This arrangement certainly has a provisional character – it simply establishes the procedure that countries are required to follow until a more definitive protocol is approved. In any case, it is worth noting that, according to this interim strategy, countries are expected to use those assessments exhibiting a better match with the targeted grades (see Figure 6.1 below). Thus, countries are allowed to use a variety of data sources, including CNAs, NAs and population-based assessments (i.e., citizen-led assessments, EGRA-EGMA, household surveys, etc.) – although, in the case of a tie (e.g., two data sources exhibiting a similar match to the targeted grade), CNAs take precedence over NAs. This is indicative of an inclusive approach oriented towards making full use of available information and providing countries with several options rather than privileging a specific data source.

---

[156] At the time of writing, to my knowledge no new strategies have been publicly proposed.

**Figure 6.1. Interim reporting strategy**
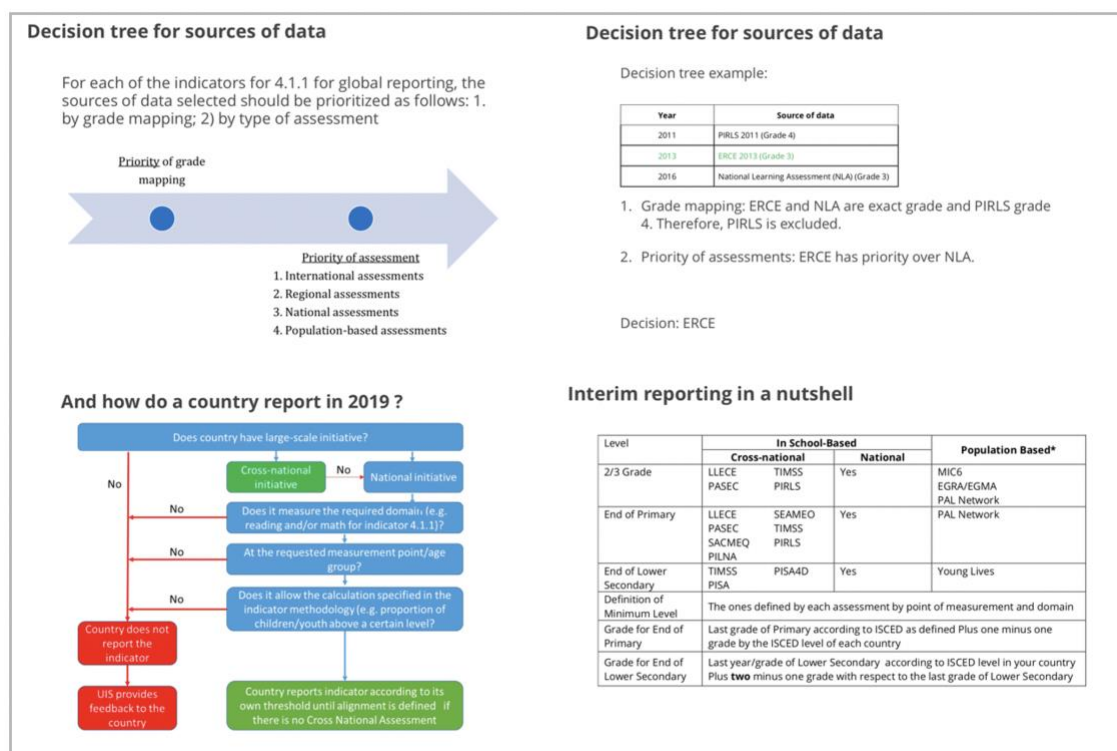


Source: Montoya (2019), *Reporting Indicator 4.1.1. Presentation delivered in the context of the 6th GAML meeting* (Yerevan, 28-28 August).

At the same time, the UIS is working on a more definitive reporting protocol that, at some undefined point in the future, should substitute the interim strategy. Although at the time of writing, this protocol is still under discussion, current efforts appear to be oriented towards maximizing data-source flexibility. Thus, and as discussed below, the UIS has encouraged the development of a wide range of harmonization strategies – most of them allowing for the use of NAs and, in some cases, even national examinations (Gustaffson, 2018; UIS, 2019a).

The remainder of the section describes the political and institutional dynamics that motivated the emergence of this hybrid approach. Particular attention is paid to the episodes in which the competing visions of what constitutes good data, as well as the tensions between the agendas of different data producers, became particularly apparent.

### 6.2.3. Veering away from a "universal test"

One of the first proposals discussed in the context of GAML was the idea of a single test. The idea was first discussed in the context of the LMTF and never enjoyed much prominence within GAML's agenda. In fact, it received limited support from the UIS, who in 2016 observed that this was only "an extreme solution" (UIS, 2016f). However, and as a consequence of the disruptive character of the proposal, the idea kept resurfacing

in the context of GAML debates, and, in many ways, shaped the expectations (and reservations) of many of GAML's regular attendees.

One of the more prominent proponents of this idea was the Center for Global Development (CGD). A group of analysts affiliated with the CGD actively promoted the advantages of what they perceived to be the most straightforward and robust approach (Birdsall, Bruns and Madan, 2016). The idea of a single test was seen as the fastest route to disseminate LSLAs even within countries that had never engaged (or shown interest) in such policy instruments. It was also a way to maximize the robustness of a global dataset, equated to *ex-ante* comparability. Importantly, the single-test agenda appeared to be driven by the belief that the potential of LSLAs lies essentially in their agenda-setting and disciplining power – or what Lockheed and Wagemaker (2013) call the "whip" function of LSLAs[157]. As captured in the words of one of the most vocal advocates of the single-test proposal in the GAML context:

> From my experience in [*name of a region – deleted for anonymity purposes*], joining PISA has been very kind of powerful in the region because it really gave policymakers something, kind of an aspirational tool and also a good political tool. So the Minister in [*name of a country*] said "We are last in PISA, we have to do better", also the Minister of [*name of a country*] said "We are last in PISA, we have to do better". And that opened the door in both cases to important reforms and more funding for education. So I thought, well, we need to get something that would be like a PISA for nine year-olds, that would have the same potential, the same political power. (PRI8)

The proposal, however, found limited resonance within the policy network brought together by the LMTF, and later, by GAML. Importantly, part of the resistance faced by this proposal stemmed from the fact that it was not perceived as being sufficiently disinterested in nature. According to a range of interviewees, it was not entirely clear to what extent some of the original promoters of the idea were acting bona fide or driven by business interests. This is not something that the present analysis was able to elucidate. However, the reservations expressed by a variety of individuals as to the "true intent" behind the single-test idea, even if fueled by mere hearsay, is revelatory in its own right – since it ultimately speaks to the governance problems that characterized the early days of GAML. Such concerns are captured in the words of an NGO staffer, who recalled that, during the tail-end of the LMTF process:

> The conversation almost bifurcated between people who have a vested interest in having the single assessment and potentially implementing it, and having the contracts for it, and those we are looking at "what does this mean for our work or others' work in a subnational or national context" […] So there's international assessment programs [*attending the GAML meetings*] and, in different times, there have been different representatives of some

---

[157] The authors draw thus a distinction between the *whip* and the *thermometer* functions of LSLAs. As whips, LSLAs *motivate* countries to undertake some form of reform in order to improve the state of education – they can thus be used by certain actors to press policymakers into taking political action. As *thermometers*, LSLA provide comparative information that provide information on key factors shaping student achievement.

of the tests that you need, there are also contractors in our meetings, you know? Pearson was in the room! And someone leaned over to me and said "Hey, you know, if one of the testing companies gets it, they'll just give it to Pearson to do it" […] So you have, you know, people who do the tests, and people who have such interest in using this test for research. (CSNGO6)

However, the ethical concerns about the commercial implications of establishing a universal test were compounded by a sense of "expert skepticism" regarding the feasibility of such an ambitious project. In reference to the meeting in which the single-test idea had first been pitched by the CGD, the above-quoted interviewee recalled that "several of us there at the table, who have tried doing such things in dozens of countries over the last decade, kind of laughed" (CSNGO6). The idea was thus largely perceived as excessively idealistic and not grounded in reality – inattentive to implementation, cost and time challenges.

Once the idea of a new universal test began to lose steam, the idea of selecting an existing CNA and extending it to new countries was vigorously pushed by some assessment agencies – the IEA and the OECD in particular. Also, and as noted above, the idea had initially been identified by the UIS as one of the most viable and pragmatic options to create global learning data. Thus, the window of opportunity opened by SDG4 monitoring needs was seized by both the IEA and the OECD as an opportunity (or incentive) to move into the development realm and expand their portfolio of countries – an endeavor already initiated through the LaNA[158] and the PISA-D projects[159], respectively. As noted by a member of the GAML-SPC:

> After we got back the question of "one test for the world", and that kind of got put to one side, then there was this question of… you know, particularly IEA was going "Well, we are already at the primary level. We're already all over the world. We should be *the test*. And then PISA, they were trying to jump in to say, "Oh well, we should be 4.1.c because we are at the end of secondary". So they were trying to occupy that space […] They were basically trying to argue, "Look, those are valid, reliable, long-standing international assessments. Why are we even debating it? These should be the tools" […] So they were, I'd say, taking an arrogant approach. And they were trying to intimidate the others based upon their technical prowess. But they failed. And they had to back down (DON6)

---

[158] IEA's Literacy and Numeracy Assessment for Developing Countries, targeted at "countries where TIMSS and PIRLS are too difficult to implement" and designed as a first step for participating in such CNAs (Mullis and Martin, 2015, unnumbered).

[159] Thus, and as documented by Addey (2016), the PISA program has progressively been extended to a number of low- and middle-income countries, and such developments have gained considerable momentum with the development of PISA-D (PISA for Development). Importantly, , the OECD has approached PISA-D as a central element in the OECD's strategy in the development realm. The author noted thus "In 2014, the OECD stated that PISA-D 'supports directly the OECD Strategy on Development and the global relations strategy of the PISA Governing Board' (OECD, 2014b, p. 19). The Organization also admits satisfaction in its ability to combine its strategy for development16 with PISA as 'the key policy instrument' in the global education agenda driven by the Sustainable Development Goals" (Addey, 2016, p. 698).

The behavior exhibited by these agencies was largely perceived as opportunistic and unhelpful. Different interviewees noted that, partly as a consequence of the prominence given to the "adopt a test" idea, the initial stage of GAML was riddled with competitive dynamics which exerted a diverting effect. A representative of a bilateral agency summarized GAML's first years of operation as follows:

> I didn't go to GAML 1 [*first meeting of GAML*], but I heard it was a bit of a crazy time. And at GAML 2, it was this sort of chaotic situation where you had all technical people talking technically to each other, therefore leaving everyone else who wasn't a psychometrician out of the conversation. But simultaneously, they were very competitive because they all wanted their test to be the top one. So you were seeing that, with IEA pushing TIMSS and PIRLS, and the OECD pushing PISA, ant the UIS trying to kind of work it out in the middle. And it just seemed it was just going to be this complete fight. (DON7)

The idea was thus met with reluctance on the part of some civil society organizations, non-profit organizations and bilateral agencies. They did not perceive it as sufficiently disinterested in nature. Additionally, there was limited appetite for such an approach within the donor community. Such reluctance was partially driven by a perceived lack of alignment with country priorities, and the reputational risks associated with imposing a specific program (or appearing to do so). As noted by the above-quoted informant:

> The problem was not a technical one. No one was questioning whether these tests were good. So, the technical and, for lack of a better word, political were getting mixed. Because the real question was, "Wait a minute. Countries have had the opportunity for many, many years to choose to be part of these tests, and for a variety of very rational reasons, they have not chosen to participate. So why would they, all of a sudden, decide this is going to be what they want to do?" […] So it was more a question of, we have to be practical about the ways and means of bringing in those countries that aren't yet participating in international or regional assessments, and how can we make this process useful for them, and not just about reporting at the global level? (DON7)

Thus, the limited headway made by this option was not only the result of a lack of buy-in within the GAML community, but also of the hesitance of many key donors who were not willing to be perceived as imposing a rigid testing program on recipient countries. Such dynamics are ultimately indicative of the limited traction enjoyed by top-down approaches to aid and development. As one think-tank research fellow put it:

> I guess that a problem with [*the idea of extending*] PISA is that PISA is funded by the OECD for the OECD [*sic.*]. So someone else would've had to put up the money for any adaptation and specifically for much lower-income countries that were not OECD members. But I still think that would have been doable, but nobody… You know, the ones that could have done this would be Silvia [*Montoya, UIS Director*], or the World Bank or the GPE. And I was talking to [*GPE officials, deleted for anonymity purposes*], told them, "grab onto this agenda". But I guess the fundamental political issue is that people don't know what the country demand for transparent, robust assessment is. And the assessment will show they are worse than other countries. And so, no one wanted to be in the position of looking like they're forcing testing on countries. So that's the work to do, to build the demand for it, which I don't think was unsurmountable, but no one has felt that's really their role on their agenda […] At some point we were talking to the Gates

Foundation to encourage them to fund learning metrics for Africa, and they also did not want to get into anything. This came from the very top, that they didn't want to get into anything that looked like they were pushing testing on countries that didn't want it. (PRI8)

### 6.2.4. The charm of a common scale

Against such a fraught and politically charged background, the UIS turned to (and, at least since 2017, expressed a consistent preference for) an ensemble-like strategy allowing for the combination of multiple data sources. According to different interviewees, the idea of a common scale, against which different learning datasets could by plotted against, gradually became attractive within UIS quarters. Not only did the idea offered a way-out of the political divisiveness caused by those approaches privileging a limited set of CNAs, but it was also perceived as better aligned with the principle of country ownership, and a reporting strategy at the service of countries' statistical needs – rather that the other way around. This vision is captured in the words of different UIS officials:

> A temporary solution for data could be that they used regional and international assessments... But the UIS has always been supporting countries. So, if we can build the capacity to carry out national assessments... and at the same time, as a by-product of this capacity, enabling them to report at the global level... then it's better than if we don't build the capacity and we just take the regional or international assessment. So, that's a key aspect: if you want to be successful, you need to build at a national level. And that's what UIS have always been favoring. (UN18)

> I know IEA will be saying that they did studies, research on the curriculum and found there are many commonalities. But the thing is, we are not looking at commonality, we are looking at the diversity, because many national assessments, they might be testing things that are not actually captured by either regional or international assessments. Also, through the SDG4, the focus is on country perspective […] So I think it's very important to take countries perspective into focus. In order to monitor we need something common, a common scale. And that's why we to develop a common scale, so that everything could be aligned to this common scale […] So it's not that we've discarded the idea [*of a single test*], but we are streamlining the idea into such a way that we are actually creating a scale for countries to map onto... for them to understand and monitor their own progress, not only the international monitoring but their own country-level monitoring. We are trying to take country level perspective into account, then it's hard to create a giant test that can fully integrate that perspective. (UN16)

To be sure, the idea dated back to the times of the LMTF – the first proposals had in fact been circulated in response to the limited appetite for more CNAs. A researcher long involved in these efforts noted:

> We determined through our consultations that there really wasn't a lot of demand for international assessments… So we actually tested kind of both ideas. One was to take TIMSS and PIRLS and just scale those up to all countries. You can imagine the people behind this [*laughter*]… the people doing TIMSS and PIRLS! So we decided, you know what? Let's try to put this issue to bed. So put it up for consultation, and we'll put it to

the task forces, and see what they think… And we all put another option, a kind of neutral international body that would kind of help with standard settings, and help improve capacity about assessment, but would not necessarily be tied to one or another. (PRI4)

Also, and as the discussion of the LMP (see section 6.1.2 in this chapter) suggests, ACER had long been a vocal proponent of the common-scale option. In order to market it, ACER strategically emphasized the value of country priorities, actionability and country ownership, and framed it as the means to circumvent or prevent turf wars among leading assessments trying to instrumentalize SDG4. An ACER analyst noted:

> The point I would like to make is, I don't think any of these processes should be about telling countries to do things that they don't see for themselves as valid to them. So, if a country sees value in participating in an international assessment fine, support them! But if a country does not, but sees participation or improvement in their capacity to undertake national assessment is of more value to them… fine, support that. If they want to do both, support that as well […] And what we are trying to do here is developing a methodology that is not imposing a particular testing program for a country. Rather we are providing mechanisms to report against SDG4 that are responsive to the choices countries make. (LASS8)

ACER thus became one of the main technical partners responsible for the development of the idea[160], known since 2017 as the UIS Universal Reporting Scales (UIS-RS). The objective of the UIS-RS was to maximize the number of countries reporting on Target 4.1 without privileging any particular assessment modality. The scales consisted of a series of learning progressions or "levels" in reading and mathematics against which the proficiency levels of existing assessments could be mapped (UIS, 2017d; 2018b).

The idea of a common scale not imposing a specific data source enjoyed the support of a variety of parties to whom in-built comparability was far from a deal-breaker, and who emphasized the need to promote an assessment strategy aligned to country needs and not exclusively oriented towards global reporting purposes and comparability. It is important to bear in mind that, both within the LMTF and the GAML community, data usability and actionability, as well as country ownership, had been prominent values. As a CSO representative put it:

> My point is, comparability doesn't really matter because, who cares? I mean, if you rank countries, what can you actually do with that information? What matters is, are there national assessment systems that engage all stakeholders, that are sample-based, that can actually show where inequalities exist? Because some inequalities are specific to the national level, you know […] So some are more focused on the national level, and some are more focused on the global comparability piece. (CSNGO4)

In coherence with this approach, many GAML regulars emphasized the importance of country priorities, and the need to support (or at least be open to the possibility of) the

---

[160] However, the project also relied on the work conducted by the UNESCO's IBE, which undertook a mapping of over 140 NAs and CNAs and their proficiency levels.

use of country-owned datasets for the purpose of global reporting. Such an approach was perceived as more in line with the spirit of the new SDG agenda and more respectful of the principle of country ownership. A country representative noted that many Member States had long been advocating to decouple certain indicators from a particular data source, and were working to convince the UIS to rely on national-assessment data in order to report on Target 4.1:

> What we see more feasible is to try to strengthen national assessments, national statistical systems, in order to provide data for this agenda. So that was our main point since the beginning. We have been trying to make the group think [*presumably, the TCG*] and consider having another effort to indue an improvement on the national statistical systems and the national assessment systems. We should consider using this information to compare countries. Of course this poses technical issues in relation to the feasibility of comparing different national assessment with different methodologies, with different metrics. We agree on that point! However, we think that efforts should be put on this side, on this direction, to strengthen the national system instead of trying to establish international projects to compare countries. (CR3)

A researcher representing a US-based research consortium emphasized the need for an approach more clearly aligned to countries' education-planning needs. The interviewee drew a distinction between the needs of "Paris, Washington or London" and those of Member States in order to capture the trade-off between a reporting strategy oriented towards satisfying the needs of the global statistical community, and one oriented towards effectively informing domestic decision-making,

> Initially everyone was very concerned about comparability. My view is that it doesn't matter if the tests are different […] They [*countries*] should use whatever assessment is most effective and efficient for them to report on these global goals and matters. It's much less about what we believe in Paris or Washington or London... than what it matters in Nairobi and Bangkok. What is important here is that they use these results to inform and act on instruction, and improve outcomes for kids […] I think at the end of the say, we need to think, what's useful for a country to improve their system maybe is not useful to report against goal 4.1. These are two separate things. And that's what people in Washington and Paris forget sometimes, they would say, "Oh, it has to be perfectly comparable, and great quality, an if not, it's worthless". But it's not about you! It's not about me! It's about children learning in Ghana, in Chile, and the US. (PRI5)

### 6.2.5. An unlikely compromise?

Despite the broad support enjoyed by the common-scale proposal, the idea was initially received with skepticism on the part of international assessment producers. In particular, the IEA and, to a lesser extent, the OECD had long emphasized the value of in-built comparability and the technical superiority of CNAs. Representatives of these organizations argued vehemently that *ex-post* comparability was a problematic idea – and that the only way of producing reliable and accurate data was through the administration of a common test. As one OECD official put it:

> To be absolutely blunt, the only way you could compare results internationally is if everybody takes the same test. There's no other way. If everybody is taking different tests, then the only way you can compare is with some kind of equating or linking study, and again, it's not easy and the results are problematic. It's highly problematic, so our view is basically that the quickest route, the most effective route, the most efficient route to measure learning in a way that is internationally comparable is to get countries to take part in international, cross-national assessments and that's it. *No other way*. (LASS3, emphasis mine)

Such emphasis on technical robustness and rigor had something of a paralyzing effect during the first stages of the GAML debate. According to different parties, the lack of progress during the early days of the Alliance came mainly from CNA's technical orthodoxy and unwillingness to compromise – that is, to relax some of the psychometric standards. A range of interviewees noted critically that many of the initial difficulties found in the production of SDG 4.1 data had to do with the disproportionate influence and power exerted by large assessment consortia, and with the IEA's immobility in particular. Representatives of the NGO and the civil society sector consistently decried the debilitating effect that such dynamics had during the first stages of the debate, and remained unconvinced by IEA's oft-repeated claims regarding the centrality of in-built comparability. In reference to the advocacy efforts of large assessment producers, and their emphasis on psychometric rigor and technical sophistication, an NGO officer observed:

> It was really just not productive. Agree with them or don't, we just kept hearing the same thing again and again. And I was… "I understand the gold standard, I get it! It's just that I have problems whether they are feasibly and relevant, and whether there's even capacity to pull it off". And then, in later conversations, there were other players that sort of weight in with these messages about feasibility and national priority. But early I the process, it was the IEA team that showed up. (CSNGO7)

Similarly, reflecting upon the reservations expressed by CNA representatives in relation to the common-scale idea, an assessment expert drew an analogy between psychometrics and religion or "priesthood", noting jokingly that "IEA, ACER, the OECD and the PISA people… all of them are like sects within the religion. So, trying to get all of them together so that they all perform the same confession and communion acts is very, very difficult" (PRI6).

In order to counter such tendencies, both ACER and the UIS emphasized the need to avoid "letting the good be the enemy of the perfect", noting that in-built comparability was not an overriding principle. Thus, the notion of "good enough data" and "fit-for-purpose data" (cf. Chapter 3) started to gain traction. The following excerpt captures this emphasis on the need for a less orthodox approach to comparability:

> I would say that global metrics in their purest sense, and by the highest standards of psychometrics, are impossible. So the question is, can we do them *well enough* that they are *fit for purpose*, right? Can we implement a model that draws upon data from countries and doesn't impose a particular testing program? That is consistent with the reporting

framework, the global metrics, but useful enough to achieve their goal… but acknowledge not to be perfect. And I fear that there will be some people who see the imperfections, and see a glass half-empty rather than a glass half-full. And so that's what I fear. And the consequence of that is to go to something like PISA, or TIMSS, or PIRLS, or you know, one of the international studies. Where a very tight testing regime is imposed. And I just don't think that's the way forward in what respects the diversity of the world. So I guess… technicians are gonna need to make compromises. (LASS8, emphasis mine)

Mm, you've probably heard of Ray Adams who is the brains behind pushing this metrics work forward. He often says, "Moving this forward would probably get me kicked out of the 'international psychometrics club'" So if you want to take this purely from an empirical perspective whether this be psychometrically valid, the answer is no. And you're never going to get it because you're never going to have a "Global PISA" […] By trying to [*inaudible - turn?*] the conversation around, we need to talk about *fit for purpose*. Is it robust enough that it allows us to have meaningful reports on student progression with the full understanding that this is not an empirical response and this is something that is not psychometrically robust in the world of psychometricians? (LASS11, emphasis mine)

At the same time, some interviewees observed that the resistance of certain CNAs was not only the product of technical considerations, but also a consequence of these organizations' fear of losing their monopolistic position in the LSLA field. Reflecting on the fact that the IEA and the OECD occupy a narrow peak in the realm of assessment largely as a result of their comparative approach[161], some observers hypothesized that the prominent position of such assessments could be negatively impacted if the idea that learning data can be compared *ex-post* was to gain further acceptance. To be sure, both ACER and the UIS put significant effort into dissipating such concerns. UIS and ACER informants noted that the fear of large data producers stemmed largely from a lack of understanding of the common-scale approach, and emphasized that such concerns were ultimately unsubstantiated. The reasoning is captured in the excerpt below:

There has been a lot of resistance because regional and international assessment, they are afraid they are being replaced, but the main thing, we are not replacing them… we are still using the regional or national or international assessments. So my understanding is the fear comes from lack of understanding. They don't fully conceptualize or understand the reporting scale… They think these insertional organizations are trying to get their items, take away their business. (UN16)

Now, I think that in the early days, groups like the IEA, and the OECD to some extent, the PASEC and SACMEQ… to some extent were threatened by the work we [*in ACER*] were putting forward. They thought their territory was being challenged, they thought that we were proposing something that would undermine their authority, their position,

---

[161] While pre-theoretical in nature, such insights are in fact consistent with the findings of specialized literature. Thus, the centrality of the comparative perspective in explaining the authority of certain organizations, and of the OECD in particular, is in fact a well-documented phenomenon. According to Martens (2007), the "comparative turn" (or "governance by comparison") was one main drivers of OECD success. Similarly, and as advanced by Grek "the OECD has created a niche as a technically highly competent agency for the development of educational indicators and comparative educational performance measures" (2009, p. 25).

the relationships, client-based, they had built with the countries participating in these assessments…. So that was quite a very strong issue and difficulty that we had to work through with the UIS […] In fact the opposite is the case. What the UIS and ourselves put forward presents an important opportunity for those existing assessment programs to gather around the table and use the mechanisms and tools that the UIS is proposing alongside… for they to use their own assessment outcomes as the data for the mechanisms that the UIS is going to need! So we would say that's an opportunity for these groups rather than a threat. (LASS10)

## 6.3. Knocking heads together: Building consensus in a fraught arena

The previous section finished with an apparent impasse as a result of the need to reconcile UIS' pragmatic considerations with the technical and political concerns of the leading data producers. It was thus noted that, while the UIS never entertained a top-down approach to data harmonization, the organization faced considerable pressure to rely on uniform measurement models. This section resumes the negotiation process – it analyzes the process through which a (fragile) consensus was crafted, and the conditions that made this agreement possible. The section therefore provides an overview of the later stage of GAML and of the UIS' efforts to accommodate the diversity of LSLAs for global reporting purposes – as a means to maximize coverage, but also to combine country ownership with technical rigor.

### 6.3.1. Striking a balance between pragmatism and technical rigor

Given the limited enthusiasm generated by the common-scale approach among the leading assessment producers, the UIS soon became aware of the need to secure the buy-in of such partners. In order to avoid reaching an impasse, a technical meeting targeting only regional and international learning assessments, along with development partners, was convened in June 2017. In the meeting, the UIS insisted upon the need to adopt a *pragmatic* approach, insisting on the "not letting the perfect be the enemy of the good" idea – a narrative long emphasized by the UIS Director. In order to get the necessary support on the part of the different partners, the UIS promoted the so-called "stepping stone" approach (Montoya, 2017). This strategy was supportive of a variety of courses of action and was crucial in securing a consensus among partners with different agendas and priorities. This approach supported three different options, each of them associated with a different temporal horizon:

- A short-term reporting strategy according to which countries are allowed to submit data of their own choice – accompanied by some basic information on the quality of assessment so that some comparisons can be drawn;

- A mid-term strategy oriented towards linking existing assessments to a common scale – building on the work conducted by ACER;

- Supporting countries to join CNAs in order to create a "critical mass" of comparable data – an objective with an undetermined horizon.

Thus, by encouraging multiple streams of work (rather than focusing exclusively on the UIS-Reporting Scales), the meeting succeeded in breaking the deadlock – with the CNAs green-lighting the pragmatic approach (Crouch and Bernard, 2017; Montoya, 2017).

However, even if the gathering was touted as a "major leap", some interviewees noted that it was the follow-up meeting, celebrated in Hamburg in September 2017, that effectively operated as a turning point. The meeting was again exclusively targeted at assessment and development partners. The relevance of the encounter lies in the fact that it was seized by the IEA as an opportunity to regain some centrality in the global reporting game – a move that signaled a certain willingness to accommodate its strategy to the mid-term approach favored by the UIS.

Thus, the IEA proposed a new linking strategy connecting international and regional assessments, an approach subsequently known as the "Rosetta Stone" or statistical linking approach (see Montoya and Hastedt, 2017; and UIS, 2017). The strategy, which placed the IEA at the center of alignment efforts necessary to put in practice the common-scale proposal[162], provided the IEA with an opportunity to vindicate the "comparability gold-standard" and to reassert its indispensability in SDG monitoring efforts. Thus, even if at the start of the GAML process the IEA had insisted on the added-value of *ex-ante* harmonization or in-built comparability, by the time of the Hamburg meeting, the organization had found a way to play a leading role in the construction of *ex-post* comparability. A GAML-SPC member described this tactical move as follows:

> So, IEA kind of went quiet for a couple of meetings because they were trying to figure out how to play the game. And then they threw on the Rosetta Stone, kind of like a rock. That would have been, I guess, the Madrid meeting. It was not even in the agenda. They just kind of threw it on the table. So that was kind of intriguing. And so then, when they come back in Hamburg, they had a proposal […] So their tactics changed, the main tactic they tried to play for with Rosetta Stone was not to rule the world, but rather to try to position PIRLS and TIMSS as kind of baseline assessment. (DON6)

Different informants have noted that the Rosetta Stone approach was, in fact, not entirely new – it had been circulating for years within the cross-national assessment community, although, until that point, none of the data producers had an incentive powerful enough to engage with the proposal. Thus, a former UIS analyst recalled that a project akin to the Rosetta Stone, oriented towards linking regional CNAs, had also been funded by the GPE in the early 2010s – but had not made progress due to a lack of sufficient funding as well as the resistance of certain regional assessments to what was perceived as a top-down approach (UN7). Similarly, a former UIS senior official noted that he had discussed a

---

[162] More specifically, the Rosetta Stone approach consisted of administering both an IEA and a regional assessment to a subset of students in two (or more) countries in each region (UIS, 2017d). Further details of this linking strategy are provided in the next sub-section.

similar idea with the IEA around 2010. At this time, one of the IEA's analysts had informally proposed this approach as a way to overcome the reticence of CNAs with regard to item-sharing (one of the arguments more frequently invoked to justify their limited effort in the generation of globally comparable learning data) (UN21). Thus, and in contrast to such "failed attempts", it is interesting to note that the (perceived) threat posed by the UIS Reporting Scale operated as a powerful incentive for certain CNAs to embrace the Rosetta Stone approach.

Thus, by the time of the GAML's 4th meeting in November 2017, the debate around the possibility of using multiple data sources and mapping them against a common scale was very much settled. The combination of existing assessments was broadly accepted as a politically and technically acceptable course of action, and a consensus had been reached on the definition of minimum proficiency levels (Montoya, 2018a, 2018b). The concept note prepared for the meeting captures the state of the debate – it was noted that the international alignment of learning outcomes was neither an impossible objective, nor one that required the imposition of a single and universal measurement process. The note concluded that "a fit-for-purpose approach to international monitoring must be achieved that supports consistency in reporting of outcomes, while being flexible enough to accommodate a variety of approaches" (p. 3).

In coherence with this all-encompassing approach, the UIS engaged in an effort to increase the availability of learning data – without privileging any source in particular. Thus, the UIS launched a series of publications, prepared by consultants and oriented towards supporting both the establishment of NAs and the participation in CNAs. For instance, the quick guide *Implementing a National Learning Assessment*, published in 2017, was conceived as a hands-on, step-by-step guide aimed to provide some basic guidance to countries willing to develop a NA. Other initiatives, including the quick guide *Making the Case for a Learning Assessment,* published in 2018, provided a number of recommendations on how to initiate a policy dialogue around the need for LSLAs – while discussing the pros and cons of NAs vis-à-vis CNAs.

Also during this time, the UIS reinforced its role as standard-setter through a series of publications identifying good practices in relation to learning assessment data collection and implementation. This particularly true of the report *Principles of Good Practices in Learning Assessment*, prepared by ACER (ACER/UIS, 2017). The document identified a number of key areas (instrument design, sampling, data management and so on) and, for each of them, identified what constituted desirable practice, as well the different strategies and steps necessary to comply with such standards. On the basis of this document, the UIS developed the *Procedural Alignment Tool* – an evaluation instrument that allowed countries to assess whether their NAs could be used for SDG 4.1 monitoring purposes in the context of the interim reporting protocol.

Finally, it is also around this time that the UIS started to actively advocate for the need to create a *global bank of test items* – that is, a repository of test constructs that would

crowdsource the items used in existing assessments. The bank of items would facilitate the establishment or strengthening of NAs, whilst allowing for the use of NA-data for global reporting purposes. In this sense, the proposal would combine concerns over country ownership with the *ex-ante* harmonization principle.

The idea has appeared in a number of UIS publications, at least since 2017 (UIS, 2017c; UIS, 2018a) although it was not only in 2019 that the proposal was discussed in some depth. Thus, a concept note published in 2019 described the rationale and main features of an SDG4 Item Ban and Exchange Platform. As per the note, the Platform was devised as a tool that would enable low- and middle-income countries to generate assessment data at a comparatively low-cost while allowing them to report on SDG 4.1.1. Particular emphasis was placed on country ownership – it was thus noted that the Platform "will allow for country ownership and will democratize testing, since countries will make their own decisions, receiving guidance rather than instructions from the UIS, which will be acting in its capacity as a facilitator" (UIS, 2019b, p. 3).

While the project has made limited headway, it is indicative of the UIS' willingness to reinforce its role as a provider of global goods, fortify its capacity-development function (by supporting countries in their efforts to strengthen national statistical systems) and, ultimately, prevent large assessment providers from exercising unrestrained power. This emphasis on such themes is clearly articulated in the words of one UIS analyst:

> [*Assessment*] it's a business. And so, because we are neutral, we can actually break the market. And in essence the UIS intention would be to develop open-sourced tools for people to be able to apply whatever they want. Also, open-source tools to assess assessments, to assess if PISA is good, SACMEQ is good? So you transfer knowledge and ownership to the population, basically. (UN18)

### 6.3.2. Bypassing the linking debate

Despite the consensus attained by the end of 2017 around the combination of multiple data-sources, a new debate emerged at around the same time. While an agreement had been reached regarding the content and the minimum proficiency levels, there was still a need to decide on an alignment strategy (i.e., a procedure to equate existing assessments to the common scale). This soon proved to be a challenging endeavor for the UIS since, once again, different organizations favored different options and insisted upon the technical and/or political superiority of their own alternative. The different strategies were developed in a series of UIS-commissioned documents and pilot studies, and presented and discussed in the context of the 5th, 6th and 7th GAML meetings. Although over this period the different options went under different names, the range of possibilities can be summarized as follows:

- The first option is known as *psychometric item-based linking* (or psychometrically-informed recalibration). It stems directly from the UIS-

Reporting Scales and has thus been promoted by ACER. This approach proposes to align existing CNAs to the UIS-RS on the basis of item-based equating, and, more importantly, to create a pool of calibrated items that can be included in any assessment (including NAs) so that *any* test can be linked to the UIS-RS (Montoya and Tay-Lim, 2018; UIS, 2017d).

- A second option is known as *psychometric test-based linking* (or recalibration through parallel tests) and builds on the Rosetta Stone proposal advanced by the IEA at the Hamburg meeting discussed above. The strategy aims to link existing regional assessments to IEA's TIMSS and PIRLS achievement scales, on the basis of concordance tables that enable the "translation" of scores across different assessments. In order to devise these concordance scales, a subset of students in two (or more) countries are administered both the IEA and the regional assessment (IEA, 2018; Montoya and Tay-Lim, 2018).

- A third option is known as *statistical recalibration of existing data* (or statistical modeling) and is based on the databases and method developed by N. Altinok, N. Angrist and H. Patrinos with the support of the World Bank (cf. Altinok, Angrist and Patrinos, 2018). The approach takes advantage of a number of countries participating in more than one CNA, referred to as "doubloon countries", on the basis of which comparable estimates of cognitive skills can be produced (Altinok, 2017; Montoya and Tay-Lim, 2018; UIS 2018c).

- A fourth and last option is the so-called *policy-linking approach* (or social moderation or pedagogically-informed recalibration). Proposed by Management Systems International (MSI), a US-based consultancy firm, this approach is the only one defined as "non-statistical" in nature, for it relies on expert, qualitative judgement as the main mechanism to link NAs and CNAs to a common proficiency scale (Montoya and Tay-Lim, 2018). Thus, the alignment of existing assessments depends ultimately on the decision made by a group of experts, who are expected to reach a consensus regarding the match of different performance levels (UIS, 2018b). This approach has been promoted by USAID, which has supported pilot studies in a range of countries along with the World Bank and DfID (UIS, 2020a).

Importantly, such strategies not only differ in terms of technical complexity and financial cost, but also in relation to more politically-sensitive issues, including their item- and data-sharing implications, their capacity to impact political and media debates, their potential to effectively inform education planning and contribute to country capacity-building and, more importantly, the possibility to integrate NAs (see Figure 6.2 below). For instance, while item-based linking is the only statistical option clearly allowing for the use of NAs for global reporting purposes, it requires to CNAs to share the items in perpetuity – an option that has been vigorously resisted by both international and regional assessments. Test-based linking, in turn, is perceived as technically robust but politically

risky, as it is not clear to what extent it could incorporate NAs at some point. Statistical recalibration is among the quickest options, but it has been noted that it lacks the necessary political buy-in since its coverage is highly limited, and the method does not allow for the use of national examinations and NAs. Finally, policy linking enjoys donor support as it is perceived as cost-effective and it holds great potential to incorporate NAs, thereby preserving country ownership. However, it is considered less rigorous and more error-prone (UIS, 2018c).

**Figure 6.2. Relationship between linking strategies and coverage of assessment type**

| | Statistical linking | | | Judgmental linking |
| --- | --- | --- | --- | --- |
| | Recalibration through parallel tests | Psychometrically-informed recalibration | Statistical recalibration of existing data | Pedagogically-informed recalibration |
| PISA, TIMSS and PIRLS | Will be used | Could be used | Yes | Yes |
| Regional cross-national assessments | Will be used | Could be used | Yes | Yes |
| National assessments | Could be used | Could be used | Not clear how | Yes |
| National examinations | – | – | Not clear how | To be used |

Source: UIS, 2019a, p. 47.

Leaving aside the technicalities of the debate, the crucial point here is that, once again, such trade-offs (and in particular the possibility of using NAs) have been strategically mobilized and exploited by the different participating organizations. Thus, ACER has repeatedly called for the need to devise a system able to incorporate NAs, highlighting the benefits of such an approach in terms of coverage, country ownership and capacity-building – and arguing in favor of a less orthodox approach to comparability. As put by a GAML-SPC member:

> Those have been the more untested methods of trying to get national assessments in. and that's, of course, where the psychometricians all freak out. Because they say, "Well, the Rosetta Stone gives you solid numbers, whereas the reporting scale had a degree of judgement and analytics in that". But the ACER argument would be that once you have the common metric and you're able to map against it, that is what gives you that comparability. (DON6)

Conversely, policy-linking and item-based lining approaches have been harshly criticized by IEA and OECD specialists, who cast doubt on the validity and reliability of this approach, and remain skeptical on the possibility of using NAs for global reporting purposes. The following excerpts express the contempt with which two analysts in these organizations perceive policy-linking and item-based linking:

> So you have the Rosetta Stone… but of course, there are also easier, cheaper ways. That's the idea of what's now called political linkage, social moderation, whatever you want to call it. But again, I think that's not adequate for trend measurement. It will lead to huge error in terms of trend measures. And I think that's not fulfilling what it's asked for. We are taking thing seriously, so our approach is probably the most conservative but

definitely the most accurate one. But when they [*ACER*] come with a cheaper approach well, if you are not a statistician, and you don't listen, that, of course, feels very interesting […] But at IEA, we don't believe in this approach. We have done some simulation studies looking at what the error margin is, and the bias. And we came to the conclusion that there's a big threat for bias. (LASS4)

What is being proposed, or what was proposed early on and still being talked about is, "Oh, we could use national assessments". I'm sorry, that ain't going to work. The national assessment in Honduras, those results cannot really be compared to the results of national assessment in Australia, that's just technically not feasible. But there are lots of people who argue that "No, you can draw up a common scale". You can take these assessment items and you can equate them and link them. Okay, theoretically that could be done, but it is highly problematic. (LASS3)

While the difficulty of reaching an expert consensus on the optimal linking strategy risked having a paralyzing effect, the UIS has nurtured again a hybrid or "hedge-betting" approach, insisting upon the fact that the alternatives are not mutually exclusive and can even reinforce one another. Thus, the UIS has supported and commissioned technical work in relation to the four options[163]. Referred to as the *portfolio approach*, this strategy emphasizes the idea that the different options "should be taken more as complementary routes than as alternative options in order to minimize risk if some of the approaches prove to be too costly, the margin of error is too high, politically-unfeasible or a combination of these issues" (UIS, 2018, p. 19). This approach is also considered more respectful of countries' priorities and context-specificities – in that provides Member States with a "menu" of options (UIS, 2018). The rationale behind this pragmatist turn is captured by a GAML-SPC member:

Silvia [*Montoya*] shifted into full pragmatics mode, and just said, "I don't have to pick a winner. So what we're going to have instead is a portfolio of options. So some countries want to do the Rosetta Stone, fine" – so she is giving something to the IEA. "And if they want to do the reporting scales", [*she said*] "fine". So in effect, she pulled rank and said, "I'm not picking any of your models because it's for the countries to decide" (DON6)

*6.3.3. The Global Indicator Framework as a rallying call*

While UIS's advocacy efforts and political maneuvering have been instrumental in advancing the pragmatic approach that appears to prevail today, it is important to bear in mind that such developments have also benefitted from forces external to the education realm. More specifically, the role played by the IAEG[164] and the UN Statistical Commission is central in explaining the receptivity of different education stakeholders to a hybrid approach that, to some extent, trades technical rigor for country ownership and data coverage.

---

[163] Note that, so far, the bulk of the work on linking strategies has focused on 4.1.1a and 4.1.1b – that is, early grades and primary education (UIS, 2018a).

[164] See Chapter 5 for an overview on the role and mandate of the IAEG.

As discussed in Chapter 5, in order to refine and improve the Global Indicator Framework, the IAEG classifies global indicators into three tiers on the basis of (*a*) their level of methodological development and (*b*) data availability. More specifically, the tiers are defined as follows:

- "*Tier 1*: Indicator is conceptually clear, has an internationally established methodology and standards are available, and data are regularly produced by countries for at least 50 per cent of countries and of the population in every region where the indicator is relevant.
- *Tier 2*: Indicator is conceptually clear, has an internationally established methodology and standards are available, but data are not regularly produced by countries.
- *Tier 3*: No internationally established methodology or standards are yet available for the indicator, but methodology/standards are being (or will be) developed or tested" (UNSC, 2020a).

In light of this classification, custodian agencies (the UIS in the case of Target 4.1) are expected to come up with work plans and refinement proposals, and to report periodically on the development of methodologies and data-collection tools. On the basis of such efforts, the tier classification is regularly reviewed and, if judged appropriate, indicators can be reclassified. In addition, such developments are expected to inform two Comprehensive Reviews of the Global Indicator Framework conducted by the IAEG – a process through which existing indicators can be adjusted, replaced or deleted (and, occasionally, new indicators can be added). The first review has already been conducted and submitted to the UNSC in March 2020[165] (UNGA, 2017).

As in the case of the indicator-development process (cf. Chapter 5), the IAEG has relied on *open consultations* as the main working modality to develop the proposal for the 2020 Comprehensive Review. More specifically, two consultations have been conducted – a first call for proposals open to any interested party, in which suggestions replacements, refinements, adjustments and additional indicators could be freely submitted; and a second consultation on the basis of a preliminary list of changes proposed by the IAEG (UNSC, 2020b). The results of this second consultation informed the 2020 Comprehensive Review proposal eventually submitted by the IAEG to the UNSC.

Given the fact that the fate of indicators depends essentially on the methodological work of custodian agencies, the periodic reviews of tier classifications (along with the Comprehensive Reviews) represent an important source of reputational pressure for the UIS. In other words, the ability of the UIS to "upgrade" existing indicators within the tier classification, and to prevent SDG4 indicators from being replaced or deleted, is perceived by the global education community and donor circles as a marker of UIS'

---

[165] The second one is expected to be conducted in 2025.

competence. In this sense, the pressure exerted by the IAEG has played a key role in fostering UIS' appetite for heterodox and innovative approaches to data production – and, at the same time, it has been seized by the Institute as a way to build consensus in the context of GAML, as well as to promote and legitimize a pragmatic (rather than idealist approach) to data production.

Two key episodes illustrate such dynamics. The first concerns Indicator 4.1.1a – that is, the learning indicator relative to the early grades. It is important to bear in mind that, whilst 4.1b and 4.1c (the learning indicators corresponding to the end of primary and the end of secondary education, respectively) were upgraded to Tier II in late 2016, two years later the learning indicator for the early years was still classified as Tier III indicator. Thus, by the late 2018, the UIS submitted a proposal to upgrade 4.1.1a to Tier II. Importantly, the proposal placed great emphasis on the potential of the policy linking approach (that is, the non-statistical and, consequently, less-rigorous approach to assessment alignment). The proposal highlighted that the policy-linking approach was one particularly in line with the country-ownership principle, in that it allowed for different languages of administration, was respective of countries curriculum, protected countries' investment in national assessments, and allowed for cultural and contextual differences (Montoya, 2018a; UIS, 2018d). At the same time, the proposal was used to create a certain sense of urgency around the need for GAML to endorse the minimum proficiency levels developed by the UIS in partnership with a wide range of CNAs – as a way to show progress and give proof of the viability of Indicator 4.1.1. Since the proposal was accepted by the IAEG in the context of its 8th meeting in November 2018, the upgrade was used to legitimize the use of the policy-linking approach in the context of GAML – noting that the use of a pragmatic approach had been key in securing the progression to Tier II (Montoya, 2018b, 2018c; UIS, 2018d).

Another illustrative episode took place in the context of the 2020 Comprehensive Review. In the open call for suggestions commented above, Eurostat submitted a proposal to replace Indicator 4.1.1 with Indicator 4.6.1 – i.e., "the proportion of the population achieving at least a minimum level of proficiency in (a) literacy and (b) numeracy skills, by sex, age and educational level" (UNSC, 2019). In response to this, the UIS Director and the Director of UNESCO's Division for Education 2030 Support and Coordination, launched a public call encouraging education stakeholders to participate in the second consultation and argue in favor of keeping Indicator 4.1.1 (Montoya and Naidoo, 2019). The UIS eventually succeeded not only in preserving Indicator 4.1.1 but in upgrading 4.1.1b and 4.1.1c to Tier I. For the purposes of this chapter, it is interesting to note that such developments ultimately proved useful for the UIS to legitimize a heterodox approach to the production of SDG4 data, and to secure consensus in a field riddled with conflict. Different interviewees noted thus that the pressure exerted by the IAEG classification had proven crucial in fostering progress – operating as a warning call or at least legitimizing a "pragmatic turn" promoted by the UIS. A think-tank analyst noted that "We are now in Tier I, which was surprising to me. I didn't think it was happening. But still, this has been productive for all the community, like an incentive to do at least

something" (PRI8). Similar views were expressed by a World Bank official – who, in reference to the Comprehensive Review, observed:

> Sometimes you need a technical justification to find a political compromise, right? The technical justification was "your indicator tier will be dropped!", [*laughter*] right? That was a technical reason, right? That, in a way, I think was critical for UIS to basically, "Okay, guys. We've been in this [*inaudible*] here for five years. We haven't been able to find the unicorn nor the highlander, so let's be pragmatic here. Because otherwise, we're all going to lose". I think those processes, again, and those mechanisms are critical to help consensus. Because if you don't have these things, people would be discussing to this day, right, because they will have no incentive to agree on anything (DON14)

## 6.4. Finding one's place: The role of the UIS in the production of global learning data

The role and position of the UIS within the realm of learning metrics have been substantially impacted by the statistical and reporting needs associated with the SDGs. This section delves into this process of organizational evolution. It analyzes how the UIS has managed to consolidate its bureaucratic influence within the learning measurement field – an area in which the Institute had long struggled to affirm its leadership and expert authority. The section comments on the internal and external conditions that made this shift possible, and follows with an overview of the circumstances that might eventually compromise the UIS' centrality – that is, those factors that place the UIS in an authoritative but ultimately fragile position. The section finishes with a comment on the World Bank's recent entry into the learning measurement field, reflecting on the possible consequences that this move could have on the position of the UIS.

### 6.4.1. UIS as an honest broker

As the previous sector suggests, when it comes to the production of global learning data, efforts under way are oriented towards avoiding a zero-sum approach and at accommodating the use of different assessments and harmonization methods. The UIS appears therefore, to have embraced a *bricolage* strategy consisting of recombining a number of already available and legitimate models, recognizing the limitations of each approach and emphasizing the potential for complementarity. Generally speaking, such a heterodox approach has been largely welcomed within the GAML community, for it has enabled the competitive dynamics to be bypassed – dynamics that in the past had hindered global efforts to the harmonization of learning data[166]. In this way, the Institute has

---

[166] It should however be noted that not all interviewees were equally enthused by the pragmatic turn pursued by the UIS. A think-tank analyst characterized such efforts as "grabbing for whatever is out, going for low-hanging fruit" and rated them "sub-optimal" (PRI8). Non-coincidentally, this view appeared to be informed by an understanding of LSLAs that privileges their agenda-

succeeded in positing itself as an honest broker driven by a public-service *ethos*, convening different parties and interests and building consensus in a fraught arena riddled with vested interests. A wide range of interviewees were appreciative of the UIS's role as mediator driven by the common interest, its efforts to create collaborative and inclusive spaces, and the emphasis placed on the principle of country ownership. Such views are captured by the following excerpts from two informants long involved in GAML:

> You cannot make progress in this work without involving organizations with high capacity. But then the question is how do you make sure than then the outputs of that do not privilege a particular organization? It's a really delicate balancing exercise... and I think the UIS has given the credentials that they are not really favoring any organization. I think they are trying to move with some people.... But they also need to satisfy certain standards in terms of how they collaborate, and what they make public, and what their agenda is. And [*it*] is not that easy. But from that point of view, I think the GAML is trying to accommodate as many players as possible. (UN17)

> [*In relation to the creation of GAML*] I think what is interesting about this conversation is knowing sort of where our limits are, in building, trusting partnerships or people who take the work forward. I mean, I don't understand three quarters of what we talk about psychometrics. Sometimes I say, "I'm going to ask a stupid question now because I don't know about this", and someone else, an academic, says "Oh, I didn't know that either", but maybe this person is a super-duper expert on child development… And someone in the World Bank says, "That's what it will look like on the ground, in reality", and I say "We did that in our micro-version, this is how we took up capacity building". So I think it's not about saying "This is the expert", it's about saying "I'm the person leading this in my organization, this is my expertise, and this is yours, and this is yours…". And altogether, hopefully, we can figure this out. (CSNGO6)

Interestingly, the Institute has managed to turn a potential liability (namely, the imperfect character of global learning data and the politicized nature of the process) into an asset – an opportunity to affirm its authority in the education measurement realm. Rather than casting the reporting process as a purely technical challenge (or emphasizing expert knowledge as the most relevant attribute of the organization), the UIS has brought to the fore the political nature of the debate as well as the vested interests that shaped it, and has tapped into its aura of publicness, neutrality and commitment to the common good as a means to bolster its credibility. Similarly, rather than addressing technical rigor and *ex-ante* comparability as supreme values, the UIS has emphasized the need to combine such principles with considerations for country ownership, and to accommodate the diversity of learning data. As a consequence, the notion of fit-for-purpose (as opposed to "perfectly accurate") data has been gaining prominence – a move that, once again, reasserts UIS' centrality. This is so the fitness-for-purpose criterion entails an element of judgement that cannot rely exclusively on technical considerations – a role for which the UIS is ideally suited on account of its aura of neutrality and public *ethos*.

---

setting potential as derived from competitive pressures – rather than their policy planning purposes. For such purposes, the appearance of robustness and perfect comparability is far more important for those understanding assessment as a policy-diagnostic tool – for whom, as noted above, global comparability is only a by-product.

Overall, the UIS has succeeded in reaffirming its centrality by exposing (rather than concealing) the messiness, complexity and political difficulties behind the production of globally comparable learning data. To put it differently, when it comes to learning data, the UIS appears to derive its authority not from an appearance of scientific objectivity or expertise, but from an emphasis on the ever-perfectibility of data, the necessarily provisional character of figures, and the political nature of the measurement debate. Such framing has proved functional in that it has enhanced the centrality of the Institute in its role as honest broker and standard-setter – rather than a mere data curator.

According to a range of interviewees, both the pragmatic change embraced by the UIS, and the success of the organization in crafting some degree of consensus in a field that has historically proven resistant to such efforts, owes much to a recent change in the UIS leadership. It was frequently noted that the leadership style of the last appointed Director, Silvia Montoya, differed significantly from the hands-off approach that characterized her predecessor. This was thus a widely shared view – even if, as discussed below, some interviewees were highly critical of such dynamics. As put by a GAML-SPC member when probed about the factors behind UIS evolution:

> To me, it's down to two words – Silvia Montoya. She's just so dynamic and has elected to bring herself into the… because it's rare, right? Someone who is statistically so solid. Typically, those people tend to be introverts and are just cranking away their data and checking and checking. And so to have a Director who has all of the technical capabilities and yet has the savvy to be able to play in the UN environment and is very forthcoming about his communications and his positioning… I put almost all of that at the feet of Silvia. (DON6)

A similar view was put forward by a former UIS analyst:

> I think Silvia Montoya is one of the key driving forces […] She has a clear goal with this, and she has taken that on board, and she's pushing this like crazy... And she has managed to get all these partners to talk, and to get to the table, and so on […] I would tend to think that the work of Silvia has been instrumental, because she's been doing a lot of work behind the scenes to discuss with the different partners, to have them come at the table, and to discuss despite different with other partners and so on. So I would honestly [*say*] that she's probably the main driving force behind developing this... and I'm not sure how thinks would continue if she were to part (UN19)

At the same time, it is important to bear in mind that the revitalization of the UIS has benefited from a series of changes in the broader environment that contributed to put the Institute on a more solid footing. First, the success of the UIS in securing a (fragile) consensus among the major assessment organizations cannot be disconnected from the fact that many of the personalities leading these organizations (whom, as noted above, were at the root of much of the inter-organizational infighting) had progressively retired. A former UIS senior official noted "We have now a cooperative space, and that's also because the old players have left, and with them, the old fights. The only one of them that remains active is Andreas [*Schleicher*], the rest are gone" (UN21). Such views were echoed by a World Bank analyst long involved in GAML effects, who observed:

Part of the early divide, or the kind of cold war between the IEA and the OECD came down to the people who were leading both of those organizations, in terms of the assessment work, not getting on with each other. But as one of those people retired and was replaced... I think there's more compatibility and collaboration between the IEA and OECD now. (DON8)

Second, the appointment of the UIS as a custodian agency for most of the global indicators (or, for that matter, indicator for Target 4.1) also played a central role in the reassertion of UIS' authority. Such assignment provided the UIS with an opportunity to claim a more prominent role in the global education architecture. Also, it enabled or legitimized the Institute to venture into measurement areas in which it had met little success in the past – as is the case of the harmonization of learning outcomes data.

### 6.4.2. A fragile position

Even if the reinvigoration of the UIS is widely recognized within the global education community, the authority acquired by the Institute in the context of SDG4 reporting efforts should not be over-stated. Thus, although there is a certain agreement that the Education 2030 mandate has contributed to ensure a much more central role for the UIS, this position appears to be fragile. This is the consequence of a range of intertwined factors. First, the limited technical expertise on LSLAs currently available within the organization has long been a matter of concern, and different interviewees noted that it could put in jeopardy both the success of the global reporting effort and the leadership or pilotage capacity of the UIS. The following observations made by the above-mentioned World Bank analyst capture such concerns:

> I don't think capacity has been built within UIS over the last few years. And I think they now have more of a skeleton staff than they did before. I think it's a very small team in Montreal. I think GAML in a way acts as an extended UIS team that helps them get some of the work done that needs to be done. In addition to hiring a technical partner like ACER to do some of the more heavy lifting. But I don't think long term capacity is being built at the UIS level [...] When I think of who is actually in Montreal and able to work on this, I can count them with one hand. So it seems kind of strange when you think of all the work that has to get done. (DON8)

Such limitations, in turn, tend to perpetuate a certain dependence on external partners – as noted in previous sections, the Institute has been drawing on external expertise for the development of most of its learning-related products. Thus, many UIS publications and initiatives launched over the last few years have been commissioned or prepared in conjunction with consultancy firms (e.g., MSI), research organizations and assessment consortia (e.g., IEA, ACER) and independent researchers and consultants. While these collaborations ensure a certain degree of sophistication in the consecution of the objectives, as well as the capacity to comply with the tight timeframe put in place by the global SDG reporting mechanisms, they are also likely to turn into a double-edged sword in the long term. It can be argued that, if not equipped with the necessary competence and

capacity of commandment, the UIS risks becoming an intermediary rather than a mediator with relevant shaping power. The "vicarious expertise" acquired by the UIS by means of partnering with others, may pose significant risks in terms of sustainability, and even of legitimacy if not accompanied by the necessary levels of public scrutiny and institutional capacity-building. As noted by different interviewees, the UIS dependence on technical partners places the Institute at risk of falling prey to special or vested interests. Remarkably, such risks did not go unnoticed to UIS staffers. A UIS informant was in fact explicit on the fact that one of the challenges faced by the UIS along such processes was the need to "tame" and the assessment industry and put it in its place. She insisted upon the fact that:

> There is a fear that the service providers will be taking over, using the issues or the technical problems to take over the political debate. So I think this is something that needs to be clarified, because a technician is only trying to provide a solution. They don't take over the discussion. The discussion is led by people who define what the issue is. This needs to be clarified. And I think there's a lot of fear and a lot of miscommunication, because they didn't understand their role. So once you understand the role of the provider, or the role of the political debate, then it will be easy to have a collaborative ethic to go forward with the agenda. (UN16)

Secondly, and as noted above, there was general agreement on the fact that the revitalized role of the UIS owes much to the ability and dedication of its current Director. While not all informants referred to such dynamics in a critical way, some observed that this represented a risk for the sustainability of the UIS' role in the learning measurement realm. At the same time, some external observers noted that it was unclear how this leadership style had been received by the UIS staff – a view echoed by an evaluation recently conducted by UNESCO's Internal Oversight Service (UNESCO, 2018), which also noted that a number of posts have been left vacant. A former UIS employee noted that "I know that a number of people have left, and the work climate is not looking good, it has deteriorated, it has really deteriorated". Likewise, a UNESCO official observed:

> While the technical lead, especially on learning assessment, has been extraordinary, the management of the Institute could have been better. And I refer to the fact that staff morale is very low… The Director has been traveling a lot and there's no second command in the Institute […] It's fantastic to have a Maradona or a Messi on the team, but we are talking about when we talk about UIS is institution- building. It's about how to build your team. (UN12)

Thirdly, some interviewees noted that the UIS's rapport with UNESCO was an uncertain one. They were unclear about the extent to which UNESCO had been supportive of the UIS' efforts in the learning assessment domain. Some observed that the UIS's efforts to gain visibility had created some friction within UNESCO. Some of the new roles assumed by the UIS overlapped to some degree with the mandate of other UNESCO units. However, different informants concurred in noting that such problems could be characterized more as "personal conflicts" and "career-driven frictions" that do not

necessarily mirror institutional differences[167]. It is also interesting to note that such intra-organizational competitive dynamics appear to be at least partially driven by financial need. A UIS former employee noted:

> We are at the beginning of a new agenda, so everybody is trying to position itself, everybody's trying to kind of creating his own identity in relation to the new agenda. And that's where there have been some... Even within the UNESCO there has been some discussion or some kind of little [*skirmish? Inaudible*] between UIS, GMR, or even UNESCO Paris […] It's basically the competition for visibility, here! It's also the competition between different individuals, which spread to teams. To me that's the biggest poise into this field... is career-driven. So, because certain individuals want to be the person who is associated with something that's game-changing. You know, these are high-ego people […] [*Different divisions within UNESCO*] they don't necessarily share all the information because to have to be added-value, so that they can claim with donors, stakeholders, that they have the added-value so they should invest there. So you have all of this environment which makes it very prone to little battles and dissention between people. But is not necessarily about disagreement, in essence. (UN18)

To be sure, the institutional distance between the UIS and UNESCO had, to some extent, proved helpful to the UIS. Thus, the prestige enjoyed by the UIS among the learning measurement realm appears to stem partially from the contrast between the UIS and UNESCO's *ethos* and organizational culture. According to some interviewees, this would have encouraged the UIS to "cultivate" this sense of distance, on the assumption that such a decoupling was productive for a variety of purposes, including fund-raising. A think-tank research fellow observed:

> Their [*UIS'*] technical caliber is quite a bit different, higher than who is in UNESCO in Paris, and the UNESCO field offices. And I think they are smart, and they know that. And they are trying to guard and protect that, and to build reputation for technical competence. And I think that Silvia [*Montoya, UIS Director*] really prizes that, and she's just trying to stay in her route, stay in her lane and maintain her relationship with her funders, Gates and other people like that. (PRI8)

However, the lack of alignment between the UIS and UNESCO's priorities appears to have had an essentially detrimental effect on the activity of the Institute – especially since, as noted by different interviewees, the UIS did not receive the financial support necessary for the challenges entailed by the production of SDG4 indicators. A UNESCO informant noted:

> I think there's a problem of strategic orientation of UNESCO. UNESCO is simply incapable of prioritizing. This is a massive area of public good where UNESCO has a comparative advantage from its position to provide. The financial situation of UIS is a

---

[167] A case in point would be the reported animosity between certain UIS and GMR figures. A UNESCO informant noted "[*GMR official – deleted for anonymity purposes*] made a very big mess, in the first GAML meeting. He made a very big intervention, questioning UIS and Silvia's credibility or legitimacy in coming up with numbers and so on... and the fact that the world should not produce a global measure and so on. But I can tell you that the GMR used the first anchored scale!" (UN18).

sign that UNESCO actually does not understand that… Because that's such a key area where UNESCO could help set standards that are objective, that are friendly for the countries, that are helping the bigger cause of education and learning for all, that is sufficiently free of private interests, sufficiently free of unintended pressures... You can name it all. So this is an area that UNESCO could have changed, but they don't. And that's sad, because it leaves the door open to another organization that may be less well-placed to guarantee minimum standards for such a process to be beneficial for the world. Of course, UIS is trying and they will get some funding for that.... but it's not the way it should be. (UN17)

This brings us to the fourth factor that places the Institute in a rather fragile position – the UIS' economic situation. The limited financial capacity of the Institute is indeed an issue consistently reported by a range of interviewees, both external and internal to the UIS. The aforementioned internal evaluation confirms that the Institute has long been in a situation of economic distress (over the period 2008-2018, the UIS' annual expenditure exceeded its revenue on five occasions), something that has forced the Institute to make some difficult decisions[168]. It is important to bear in mind that, as a UNESCO Category 1 institute, the UIS is funded through *voluntary* contributions from UNESCO and from external donors, and that UNESCO's share has not been particularly stable (oscillating between 30% and 46% during the period 2008-2018) – a situation compounded by the withdrawal of the World Bank funding in 2015. Since then, GPE remains the only multilateral supporting the UIS. The bulk of the external funding comes therefore from bilateral organizations (with Canada as a host country, remaining the most stable contributor, and with important contributions from Sweden, Norway, the UK and Australia). Private entities such as the Hewlett and the Bill & Melinda Gates Foundation have also made significant contributions over the years (UIS, 2018).

Remarkably, UIS's work on learning outcomes stands out as one of the areas in which the UIS has eventually[169] been more successful in terms of resource mobilization. In 2019, the UIS secured funding from the Bill & Melinda Gates Foundation specifically focused on the measurement of learning outcomes (Montoya, Beeharry and Woolf, 2019).

---

[168] For instance, the UIS staff was downsized by one third in 2017 in light of its financial situation. Also, the UNESCO evaluation has suggested different options to reform the current funding model – including a number of user-fee based models in which users pay for UIS data. This would entail a shift from a "data as a public good" model to a "data as a semi-public good model" (UNESCO, 2018).

[169] It must be noted that this is a recent success – a number if interviewees recalled that, during the early stages of GAML, the production of SDG4 did not received enough support on the part of the donor community. Early in 2017, a GPE official noted "I don't see people with financing moving together. The UIS has the mandate for learning outcomes, but they only have investment for the usual work. And it's a far more complex thing. The concern is that I don't see donors mobilizing for that. That's interest in the discussion but I don't see the financing coming in, and we are on 2017… things are moving a little bit slow […] So there are the political challenges and the financial challenges. And they are linked! Because if you don't finance capacity... There's probably a lack of investment in what it takes to get there, to get to the universal scale linking tests" (DON9).

The UIS is also one of the main beneficiaries of the DfID's BESTA[170] program – with the allocation of £4.6 million earmarked for the production global data, tools, and methods available to measure learning (DfID, 2018). At the same time, such reliance on donor funding and the absence of a more stable budgetary basis, contribute to deprive the UIS from the necessary autonomy to develop their own agenda (since external assistance is typically tied to specific projects) and perpetuates the Institute's dependence on consultants to fill knowledge gaps or even to perform core tasks.

Lastly, the very governance of GAML represents a fifth factor exerting a detrimental effect on UIS' authority. While the UIS partially derives its authority from its commitment to transparency, inclusion and a participatory and democratic approach to decision-making, the cost of such principles appears to be progressively higher. With an increasingly large and diverse membership, GAML risks becoming an inoperative space. While the UIS has been highly effective as a broker mediating between data producers, the inclusion of country representatives in consensus-building exercises is proving challenging. A World Bank representative noted:

> We're at the stage where, in order for those tools to have legitimacy and to be scaled up, you have to bring in countries. You have to start having that conversation, asking countries to pilot these tools. At the same time there are problems ahead, or challenges that need to be faced in terms of how GAML brings together people from the country level, because what we found at our last meeting in Hamburg was that… we had something like, I think maybe 20 or 30 countries, which was great, but the people who attend these more technical meetings are not necessarily people who have decision-making authority […] So there's still that issue of how do you bring in countries-- invite countries to these GAML meetings, bring the right people in who can comment on the technical side, but also ensure that the people in these countries who have the power to make decisions and say, "Yes, we agree with this. Yes, we will follow these protocols" (DON8)

In addition, the need for the UIS to demonstrate progress, as well as to comply with donors' priorities and timeframes, makes it increasingly difficult for GAML to rely on consensus-building as the main decision-making modality. In fact, and as noted in the previous section, the UIS has resorted to GAML's *technical* meetings (i.e., meetings bringing together donors and data producers) as a means to overcome political impasses. The reliance on such agreements is however difficult to reconcile with the emphasis placed on the plenary meetings and participatory procedures – thus, a certain tension emerged between the negotiations of the boardroom and the reliance on "backroom deals".

In the light of such dynamics, some civil society representatives noted that, despite the emphasis placed on deliberative forums (such as the GAML's plenary meetings or the task forces), the UIS frequently steered the debate so as to secure specific outcomes. In other words, some perceived the GAML to operate as a rubber-stamping board in which

---

[170] Acronym for Better *Education Statistics and global Action to improve learning*.

the UIS was simply looking for the "seal of approval" of previously negotiated agreements. It was also suggested that the inclusion of country representatives followed essentially a tokenistic logic. An interviewee argued that, even after the establishment of the SPC, much was being discussed behind closed doors:

> I guess the biggest impact in some ways, the establishment of the SPC, I think the fact that that was established was an acknowledgment that the current practice, the current way of working was not effective and it was problematic in many ways. So I think that was acknowledged and there was in a way broad agreement on the process being problematic. When it comes to the more, what to call it, the actual work of the GAML, and all the criticism we have raised there, I don't think we've had much if any results to be very honest. And I think it goes to show how, in a way, the GAML is just a rubber stamp for things that would have happened anyway in a way. (CSNGO5)

Asserting the veracity of such impressions is beyond the scope of this research. However, it is worth noting that such observations are consistent with the fact that, when asked about the rationale behind certain GAML decisions, the knowledge of many interviewees appears to derive from guesswork and rumors. The fact that agendas are decided in-between meetings, and that operational plans changed frequently without going through GAML's plenary, reinforced such perceptions. Also, GAML's plenary meetings appeared to be heavily scripted – with attendants being presented with a narrow set of options rather than invited to propose alternative routes. Finally, references to what happened "behind the scenes" were recurrent in the interviews. Ultimately, such dynamics are indicative of the fact that, while the democratic imperative and the inclusion expectations placed on GAML are key sources of legitimacy for the Institute, if taken to extremes they can also operate as constraining elements that might hamper the UIS' ability to deliver its mandate – thus creating a perverse incentive for the UIS to preserve a formally democratic structure while moving the real *locus* of decision-making away from these democratic and participatory spaces.


*6.4.3. Seeing double? Competing approaches to the production of global learning data*

It is important to bear in mind that, even if the UIS saw its centrality considerably enhanced as a result of the advent of SDG4, its authority is permanently negotiated and ultimately depends on the position and role of other IOs. In this sense, it is worth noting that the UIS is not alone in its efforts to create globally-comparable learning data. Recently, the World Bank undertook a series of initiatives similarly oriented towards the production of global learning figures. While the ultimate impact of such a move over UIS's authority remains to be seen (for many of the initiatives are in their infancy), an overview of such developments is important to better understand the unstable position of the UIS in the realm of education metrics.

As discussed in the previous chapter, the World Bank has placed a growing emphasis on the need to improve learning conditions and outcomes – a discursive turn that culminated

in the release of the *World Development Report 2018, Learning to Realize Education's Promise* and, later on, the report *Learning Poverty: What Will It Take?* Such publications insisted on the need to prioritize learning and shared a number of recurring themes. These included the need to draw a distinction between schooling and learning, and the risks entailed by the so-called global learning crisis, understood in turn as the product of low-quality education systems (cf. World Bank, 2017; World Bank, 2019a).

Consistent with such a narrative, learning measurement has been gaining prominence within the World Bank priorities, including analytical and operational work. The growing emphasis on learning poverty has for instance informed the World Bank Group's Human Capital Project (HCP) – a project conceived as an awareness-raising effort and intended to "build political commitment for reforms and investments in human capabilities through advocacy, measurement, and analytical work" (World Bank, 2019b, p. 6). One of the flagship products of the HCP has been the Human Capital Index (HCI), a new macro indicator launched in 2018. The HCI is a composite index combining metrics relative to different dimensions of human capital – specifically, survival from birth to school age, health, and education. The education component is in turn measured through a new indicator – the Learning-Adjusted Years of Schooling (LAYS). The measure aims to capture education quantity and quality. The latter is measured on the basis of Harmonized Learning Outcomes (HLO) – an indicator combining multiple testing programs (World Bank, 2020). The methodology used to harmonize scores relies on the work of N. Angrist, H. Patrinos and N. Altinok – and is indeed one of the linking strategies considered by the UIS for the production of SDG 4.1 (Altinok et al., 2018).

The promulgation of the Learning Target (LT) also needs to be understood as part of such efforts. Launched in October 2019 in the context of the World Bank/IMF Annual meetings, it aims to draw greater attention to the learning crisis in order to galvanize political and technical action (World Bank, 2019a). The LT has an uneasy fit with the Education 2030 agenda and risks having a "crowding-out effect"– operating in practice as an instance of "counter-targeting". While the overlapping and sometimes conflicting relationship between SDG4 and the LT has been discussed in Chapter 5, the relevance of the project for the purposes of this chapter lies in the fact that it has consolidated the World Bank's engagement in the production of globally comparable data. This is so that the Learning Target has a Learning Poverty Indicator associated to it. The new indicator combines measurements relative to school access and learning into a single figure, developed with the UIS and closely aligned to LAYS[171]. Very much like the HCP, the Learning Poverty Indicator is conceived as an awareness-raising and an "early-warning indicator" oriented towards accelerating progress tools (see World Bank, 2019a, p. 7) rather than a policy-planning or diagnostic tool. The logic behind the indicator is captured in the words of one of the World Bank analysts:

---

[171] More specifically, it can be defined as the "weighted average of the share of the population below the minimum proficiency level adjusted by the out-of-school population" (World Bank, 2019b, p. 7).

I would like countries to pay as much attention to learning and to be as embarrassed by the results in terms of learning poverty as they are for monitoring the poverty, right? There is no president that I know of, no country that is proud or that doesn't feel embarrassed by the results in terms of monitored poverty, right? There are ministers that get fired because poverty numbers go up, right? […] So with learning poverty, now, you have the president of the country asking, "What is my human capital index, right? What is the learning poverty?" I think one of the attractiveness of the human capital index is precisely because this is something that ministers of finance understand, presidents understand. And it's a lot more sexy, if you will, right, than, "How much is my learning," or "My average learning went from 300 to 308 or 296". I mean, what is that? Nobody knows what that is! That's a complete abstraction! That's I think also one of the reasons why I think learning poverty is attractive, right, because of that simplicity of the narrative. (DON14)

In addition to the Learning Poverty Indicator, in 2019 the World Bank launched a number of initiatives similarly focused on improving the global comparability of assessment data172, including the production of the Global Learning Assessment Database (GLAD), a microdata dataset storing and standardizing data from a variety of CNAs and NLAs. Such efforts are part of the Global Comparability of Learning Outcomes (GCLO) component of the Systems Approach for Better Education Results Umbrella Facility, SABER-UF, an initiative funded by DfID and the Bill & Melinda Gates Foundation.

Such efforts are in contrast with the World Bank's limited engagement with GAML. In fact, in the debate around SDG4 indicators, the role of the World Bank has long been something of a question mark to multiple interviewees. Thus, only one World Bank analyst, Marguerite Clarke, has been consistently attending and following GAML meetings. While this official was in fact highly engaged with GAML efforts (chairing the task forces on Indicator 4.1.1 at one point), to some GAML regulars it was unclear to what extent such commitment reflected an organizational priority. The fact that, since Clarke is no longer based at the headquarters, and that representatives from the Education Global Practice or the statistical services rarely showed up at the meetings, reinforced such doubts. In addition, some noted that it was unclear to what extent a single analyst was able to influence the World Bank education priorities, or to what extent other education specialists or task team leaders were sufficiently paying attention to her input. An interviewee noted, "She's very good, really lucid, but the Bank is a big machine that simply does not think" (UN21). A researcher affiliated with a US-based development institute noted:

---

[172] Note that initiatives oriented towards improving the comparability of learning data are complemented by efforts towards the improvement of the availability and quality of data. One of the most relevant initiatives in this area include the launching of the Learning Assessment Platform (LeAP) – an initiative to support assessment-related initiatives in client countries, supported by Russian development, together with the inclusion of a new indicator on IDA support capturing "the number of Bank-supported completed large-scale assessment rounds at the primary or secondary levels" (World Bank, 2019b).

The World Bank… they really lack the expertise. Marguerite is the single exception. But a single person cannot bear the burden of an institution. Also, I think there was some naivete, they were naïf in thinking that you could get all the psychometric organizations to collaborate. I think they underestimated what did it take […] Also, over the last ten years, the Bank has been losing technical expertise in education. Economist have prevailed, but skilled people, people with an expertise on infrastructure, on textbooks, on teacher training… these people retired and have not been replaced. Since 2010, they have become weaker, they were left behind, also considering is not like they have the mandate. So then the only person knowledgeable about assessment was Marguerite. (PRI6)

Recent developments, including the HCI and the GCLO efforts, suggest that, despite the World Bank's limited engagement with GAML, the production of global learning data has gained considerable prominence in its organizational agenda. Such efforts have been perceived by many as an instance of mission creep – an attempt to expand into new domains. Some UNESCO-affiliated interviewees were highly critical of the HCI project given its overlap with UNESCO's mandate, and pointed to the risk of duplication[173]:

I'm quite concerned. So to me, this [*HCI*] didn't bode well for the Bank and its role. It really came across as a pretty half-baked idea. What's the sustainability here? Who's all this for really? […] It doesn't seem to be in much of a leadership role, and my worry is it's like other development agencies, they are more concerned about having a branded product than actually having an impact. (UN9)

Interestingly, the World Bank was not oblivious to such risks and instituted a preemptive strategy to dissipate such fears. Thus, a special effort has been made to emphasize the alignment of the UIS and the World Bank work. The observations of one of the analysts in charge of the HCI are indicative of an important degree of reflectivity on the reputational risk associated with organizational overlap:

I think externally, there's always an issue of how you make sure that you're not perceived as a threat to the mandate of any other agency. How you make sure the people understand that what you're trying to do is to add, not to take away from anything for anybody, right? […] And part of what we definitely wanted to do is to make sure that with UIS, with whom we had a very good relationship… we wanted to make sure that relationship was preserved, respected, strengthened, precisely to help, play the proper incentives for everyone […] I think there was a deliberate intention from our side to make sure that we minimize the space for any noise of this nature, I mean, to the point that in all of our communication, we try to be as clear as possible, that the our definition of minimal proficiency level is actually the SDG 4.1.1, in line with the GAML process. So we're not coming with that out of our own mind. We present the learning, and we want the Learning Poverty measure to be seen as a joint product from UNESCO, UIS, and the World Bank […] It was very deliberate and very explicit or intention to make sure that this was actually a joint effort, right? I mean, we had countless meetings with UIS. In the process

---

[173] The nature of the work of the Bank was also criticized on technical grounds. A former UIS official noted, "Really, they do things the wrong way. They do! They get things wrong, and they create so much noise. Certain economists, not all of them, but certain economists… it's noise. They pull out a number which they need to use as a dependent variable, and when you tell them, 'it's not one number, there are 44 numbers', they get crazy, they don't know what to do. Some of them, not all of them" (UN21).

of producing its numbers, we made sure that they were… we benefited a lot from their knowledge, and we wanted to make sure that they were aware of every single step of the process and when we [*inaudible – launch?*] the numbers, we [*inaudible – launch?*] it as a joint product between the World Bank and UIS. (DON14)

In order to reinforce this image of inter-organizational alignment and coherence, in 2019 the World Bank and the UIS signed a Memorandum of Understanding (MoU). According to World Bank informants[174], the MoU had a twofold objective. On the one hand, it was devised as a means to show clarity of purpose and demonstrate a cooperative (rather than competing) attitude – both to the global education community and to donor circles. As noted by one of the specialists in the Education Global Practice:

> I think one of the reasons why this partnership was sort of formalized in this MoU was to send a message to the global community that we're working on this together and that we're driving this towards the same goals. So before we were working sort of… we were always collaborating, but there wasn't this one message externally, UIS. And the other thing is we're sort of leveraging each other's strengths in this way. UIS has a particular strength. The Bank has a particular strength, comparative advantage. And so we're formally bringing them together in efforts to drive certain global objectives. (DON13)

On the other hand, the MoU was also conceived as a means to give some stability to the UIS/World Bank collaboration – making it less dependent on the goodwill and positive attitude of the officials in charge, and less vulnerable to changes in organizational agendas.

> It's an expression of interest, and it's completely non-binding, okay? And there's no money involved, right? So it's one of those things that it forces a conversation between parties, right, and helps clarify expectations […] And actually, a lot of times, often it's more useful internally than externally, right? Because our own bureaucracies have priorities, right, and priorities change. So having this MoU creates a mechanism for compliance that we can use within our own bureaucracies, right? So I believe that these things are actually, potentially, more useful internally, as ironic as it may sound, right? Because the reason that a lot of times that these relationships become dysfunctional, I think, is because, again, right, bureaucracies move and shift and priorities change. (DON14)

On the basis of the MoU, and encouraged by some key donors (cf. DfID, 2019), the World Bank and the UIS have recently started to collaborate on a number of initiatives, including policy-linking efforts (cf. section 6.3.2 above). It is also important to bear in mind that, beyond the MoU, there are efforts under way to establish an *interagency initiative for closing the global learning measurement gap in internationally comparable large-scale assessments.* As per the World Bank note, the initiative aims "to ensure that, by 2030, countries' student assessment systems can produce regular and comparable data for early primary and end-of-primary grade learning, to enable policy makers to develop evidence-

---

[174] Given the recent character of such developments, along with difficulties to gain access to certain UIS informa, it was not possible to inquire into the UIS expectations rationale behind the partnership.

based policies to improve teaching and learning, to ensure reporting on SDG 4.1.1 with an initial focus on SDG 4.1.1 (a) and (b), and to help benchmark Learning Poverty indicators" (World Bank, 2019b, p. 18)[175].

The impact of the UIS/World Bank collaboration will only be seen in the long-term – at the time of writing, the partnership still appears to be in the making. In any case, the relevance of such an alliance is indicative of a certain tension between differentiation and collaboration imperatives faced by IOs – in the sense advanced Kranke (2017a). Thus, while both organizations need to signal a cooperative attitude and demonstrate their ability to work closely, they also need to differentiate their products as a means to either extend or guard their own turf. Such competing pressures might have a particularly detrimental effect on resource-constrained organizations such as UIS – in that they create an incentive for expanding and diversifying the organizational portfolio as a means to remain relevant. Overall, the UIS/World Bank partnership provides further evidence that the learning assessment field is far from settled, and it retains an unstable quality that warrants attention in the future.

**Wrapping up**

This chapter has provided an overview of the organizational dynamics and micro-processes shaping the assembly of Indicator 4.1.1. It has shown that, despite the consensus on the indispensability of global learning data, the production of such data represents a challenging endeavor. The difficulties encountered in the process appear to be the product of two analytically distinct dynamics. First, the growing centrality of assessment data is perceived by a range of data producers and data collectors as an opportunity to consolidate and expand their portfolio of activities and areas of influence – a process that inevitably creates issues of rivalry and overlap, sometimes exacerbating (or reviving) decade-long conflicts. Second, the collective nature of such an ambitious measurement process has laid bare the fact that multiple (and contradictory) expectations are placed on the use of both LSLAs and global monitoring efforts. As noted by a World Bank informant, "All measures are wrong and imperfect. But some of them might be useful" (DON14). The question that naturally arises is to whom and for what purpose they are useful – and it appears that different organizations have found different answers. The tension between technical rigor and country relevance has thus been an undying one. While the UIS has managed to strike a balance between these competing priorities by promoting the merits of *fit-for-purpose data*, the viability and costs of such a heterodox approach are still to be determined.

---

[175] Interestingly, it is on the basis of this initiative that the bank-of-items proposal advanced by the UIS has received some impetus. Thus, one of the expected outcomes of the UIS/World Bank interagency initiative is precisely the development of an item-bank platform that could be either added to existing NAs or to create new assessments (World Bank, 2019b).

# Chapter 7

## Conclusions

This final chapter synthesizes the main findings of the dissertation and discusses them in relation to the research puzzles that the thesis aims to address. To this purpose, the chapter outlines first the empirical insight gained from this thesis. It then connects these findings with broader academic debates and discusses the main implications of the dissertation for two separate research agendas, namely, the practice of norm-setting and the production of global data. The chapter continues with a reflection on the practice of consensus-making in polycentric environments (and the conceptual and theoretical challenges posed by this theme). The chapter closes by delineating avenues for future research and further lines of inquiry.

### 7.1. Recapitulating

This dissertation aims to examine processes of global goal-setting in education through a case study of the negotiation of Sustainable Development Goal 4. While much of the public comment and scholarship on global goals has tended to focus on the relative impact and effectiveness (or the lack thereof) of such normative instruments, this thesis takes a step back and focuses instead on the very process through which such goals came into being. Thus, this research seeks to re-embed the SDG4 targets and indicators into their process of production. The objective is to shed some light on the architecture for deliberation that underpinned the negotiation of the SDG4 targets and indicators, and to unearth the arguments that were invoked during their construction, together with the conflicts eventually obscured in the final wording. The thesis explores such questions by concentrating on the role of international organizations in these processes – a comparatively under-discussed theme in the early accounts of the making of the SDGs. The thesis attempts to bring into focus the impact of organizational routines and incentives, as well as the dynamics of inter-organizational interaction mediating in the construction of normative frameworks. The overarching research objective is:

> *To examine the negotiation and monitoring of Sustainable Development*
> *Goal 4 by bringing to the center of the analysis the culture and interests*
> *of the international organizations partaking in these endeavors, as well*
> *as the institutional architecture in which inter-organizational interaction*
> *took place.*

In order to address this objective in a historically-informed and context-situated way, the dissertation engages first in a historical overview of the use of international commitments and aspirational targets in the education field. Chapter 2 offers a retrospective account of global goal-setting in the area of education – ranging from the early experiences around

the notion of Universal Primary Education during the 1960s, to the advent of Education for All in 1990. The chapter highlights how, in the realm of education-for-development, the reliance on aspirational devices exhibited both an enduring quality and a mutating nature. It is noted that, while the practice of global goal-setting has become progressively institutionalized and routinized, it has retained an open-ended nature – with each of the global encounters, declarations or pledges putting forward a number of procedural innovations and facing specific challenges. The chapter also draws attention to the constitutive character of goal-setting experiences which operated as field-configuring events. Thus, episodes such as the creation of Education for All in 1990, or the Dakar EFA Goals adopted in 2000, contributed crucially to reorganize the global architecture of education-for-development, leading to the emergence of new actors and governance spaces, as well as to the redefinition of the mandate of existing organizations. Finally, the chapter alludes to the emergence of two parallel agendas, associated with different communities of practice – namely, the MDGs and the EFA goals. Whilst at the moment of the adoption of the Millennium Declaration this overlap was the subject of limited comment, such developments ended up shaping the negotiation of the SDG4 agenda in a decisive way.

In the following chapter I set up the theoretical framework orienting the research in order to articulate the theoretical interface between global goals, quantification practices, and international organizations as objects of study. It is first argued that, while global goals constitute an elusive subject of study within IR scholarship, the practice of goal-setting can be analyzed as a vehicle for normative change – that is, an instance of norm-building, as advanced by Fukuda-Parr and Hulme (2011). I argue that the thesis could benefit from (and potentially contribute to) a literature concerned with the role of IOs as creators of norms rather than mere transmitters of norms created elsewhere. I identify a number of variables exhibiting explanatory power and analytical value when it comes to understanding processes of normative change within IOs. This includes the role of institutional sub-cultures (Nelson and Weave, 2106), expert authority (Broome and Seabrooke, 2012) or organizational leaders (Hall and Wood, 2018).

Chapter 3 also articulates global goal-setting practices as quantification exercises. Following Fukuda-Parr and Yamin (2013), it is argued that global goal-setting practices not only have a normative effect, but are also associated with knowledge effects since they have a direct impact on the very conceptualization of development issues. On the basis of such observations, I discuss the need to examine in a more systematic way the constellation of organizations engaged in the production of globally comparable education indicators, as well as the effect of indicators on their own creators and producers. The chapter outlines a number of analytical categories to bring greater conceptual clarity to the study of such issues, including the four-phase trajectory of indicators advanced by Davis, Kingsbury and Merry (2015), Rocha de Siqueiras' notion of "good enough data" (2017), and Freisten's (2016) approach to indicators' organizational effects.

Finally, the notion of field theory (Adler-Nissen, 2013; Cohen, 2018) is introduced as a mode of inquiry particularly apposite to incorporate a relational perspective to the study of the norm-building and quantification processes triggered by SDG4. While literature on the *internal* organizational dynamics behind such processes has proliferated over the last years, it remains less clear how different organizations sharing a mandate on a given issue-area negotiate and build consensus during processes of international norm-building or norm-making. As advanced by the Bourdieusian tradition, a field approach brings to the fore the question of inter-organization interplay and organizational interdependence. Hence, a field *lens* allows me to avoid an actor-centric approach and to incorporate a relational perspective, as well as to capture the specificities of transnational governance processes characterized by fuzzy hierarchy lines – as has been the case with negotiation and monitoring of SDG4.

The methodological approach orienting this research is presented in Chapter 4. The chapter describes first the data-gathering strategy employed for the compilation of the primary source of evidence informing the research – namely, 99 semi-structured interviews. The chapter thus describes the specificities of two non-probability sampling approaches used to recruit the respondents (i.e., snowballing and purposive sampling), as well as the challenges associated to a context of elite interviewing – in particular, the difficulty to determine a saturation point. In the chapter I also elaborate into the rationale behind the two interview guides on which the research relied, and reflect on the practice of online interviewing – identifying the opportunities and challenges imposed by this modality, as well as the strategies I employed to mitigate the limitations derived from by a computer-based medium. The chapter also discusses briefly two additional sources of information that were used to triangulate the insights gained with the interviews – namely, documentary data and non-participant observation. In terms of analytic strategies, I identify a case-centric variant of process tracing (Beach and Pedersen, 2013) as a means to make sense and interpretate the empirical data collected for the purposes of the research. The chapter discusses also how the interview-coding strategy was informed by this specific analytical strategy and the objectives of the research, and finishes with a note on the strategy used to present the empirical findings of the research – namely, thick description.

The main empirical findings and insight gained by the thesis are discussed in Chapters 5 and 6. The former grapples with the first objective of the dissertation, namely, *to understand how the normative statements and associated targets encapsulated in SDG4 are agreed upon and negotiated within and between those international organizations with a mandate or agenda on the global governance of education*. To this purpose, the chapter reconstructs and analyzes the process through which SDG4 came into being, from two distinct but complementary entry points that correspond to two separate sources of conflict and tension.

One entry point concerns the negotiation over the very architecture of the debate – that is, the process through which a consensus was reached regarding the legitimate *locus* of

power, decision-making structures and deliberation styles. This proved a particularly challenging enterprise given that, since the turn of the century, the global education agenda had been informed by two separate sets of goals (EFA and MDGs) associated with different decision-making architectures and with different communities of practice. The negotiation of SDG4 thus required finding common ground regarding the venues, working modalities and negotiation practices deemed legitimate for global- and target-setting purposes. International organizations and other development partners engaged in the post-2015 process with the aim of preserving their relative positions within the global education architecture – that is, in order to secure a position of centrality. The negotiation styles and consensus-building scripts favored by different organizations were therefore in agreement with their organizational culture and institutional legacies. UNESCO, on the one hand, favored a bureaucratic style relying on regional and constituency-specific consultations. In the context of these consultations, UNESCO played a consolidating role through its secretarial capacity. Specifically, UNESCO tapped into its convening power to secure organizational visibility, and was adamant ono the need to reinvigorate he EFA machinery as the primary space for debate. The preference for formalization and heavily-scripted negotiations was also espoused by different NGOs and CSOs, and especially by those more heavily and directly involved with the EFA Steering Committee. Somewhat paradoxically, these organizations featured among the most ardent rule-followers and those more critical of other organizations "breaking protocol" and favored informal engagement modalities and ad hoc platforms, or prioritized the UN venues. Conversely, other organizations felt more at ease with informal channels and loosely-structured negotiation venues. In fact, organizations such as Brookings proactively resorted to forum-making practices, thus creating parallel or additional debate venues. Also, it is remarkable that those organizations and donors with a more tenuous or recent rapport with the EFA community proved able to play the UN rooms to their advantage. In addition, they were more likely to perceive the EFA machinery as a bureaucratic apparatus playing an obtrusive role and stifling innovation. Overall, the coexistence of two workstreams proved to be one of the most persistent sources of friction – a development that suggests that the participatory and polycentric nature of the post-2015 debate cannot be mechanically equated to an absence of conflict or hierarchy, or to the suppression of power-laden dynamics.

A second entry point to the negotiation of SDG4 targets concerns the content of the agenda. The post-2015 debate became a site of normative struggle around the relative priority given to the improvement of learning outcomes. The negotiation of SDG4 was instrumental in the rise of the learning agenda and the consolidation of a learning turn within the development field. It was noted however, that whilst the negotiation of SDG4 operated as a precipitating factor behind such a normative shift, the learning turn has also benefited also from a number of predisposing factors relative to the advocacy efforts and analytic work conducted by actors that were not necessarily prominent in the post-2015 debate. It was also highlighted that efforts to culminate and consolidate the learning turn as a normative shift did not go unchallenged. The open and participatory nature of the post-2015 process became an opportunity for an alternative normative community to rally

around the idea of a quality agenda. Remarkably, this group of antipreneurs instrumentalized the post-2015 process as a *show of force* – certain NGOs and CSOs perceived the learning-outcomes debate as a Litmus test for their relative influence and centrality within the education-for-development realm. In other words, what was at stake was not only the relative importance of learning outcomes vis-à-vis other education priorities (or proxies of quality), but also the authority and centrality of CSO constituencies. As a consequence, the debate around the question of learning outcomes progressively acquired a polarizing quality. The drift between those favoring an outcome-centered approach and those concerned with the narrowing effect of such a move devolved into a seemingly irreparable schism. The struggle was eventually solved by recourse to an ambiguous wording and an explicit emphasis to dimensions of quality other than learning outcomes. However, because this final agreement did not rest upon a process of ideational convergence or rapprochement, tensions over the relative priority deserved by the improvement of learning outcomes, and the use of learning outcomes as a proxy for education quality, continued to resurface in other forums and streams of the debate (most notably, the construction of the SDG indicator framework). At the same time, I argue that the absence of a genuine consensus around such questions risks having a fragmenting effect – leading discontent agencies to place limited emphasis on the SDG4, or to reinterpret the framework according to their own priorities.

Chapter 6, in turn, is concerned with the second research objective of the dissertation, namely, *to understand how the quantification needs brought about by the new agenda have impacted on the institutional agendas of the organizations involved in the production of SDG4 indicators, as well as on the relationships between data suppliers and data harmonizers.* I focus on the organizational dynamics behind the production of Indicator 4.1.1, relative to the monitoring of learning outcomes, in order to gain insight into the mechanics of global quantification, i.e., the process through which indicators are coupled to specific data suppliers, statistical routines and reporting standards.

The chapter highlights how the production of Indicator 4.1.1 has entailed the emergence of new forms of interdependency between the UIS and assessment producers. Whilst, historically, the UIS has relied almost exclusively on administrative data collected by national statistical offices and Ministries of Education, the need to report on learning indicators has institutionalized the relationship between the Institute and a new kind of data provider – namely, assessment producers. This has not been without challenges since assessment producers have frequently approached the production of SDG4 data as an opportunity to consolidate or even expand their visibility, prestige and outreach. Thus, the UIS has been faced with the task of striking a balance between the (sometimes conflicting) interests of different assessment producers. More importantly, the UIS is also responsible for ensuring that the SDG4 data-collection protocol (and the assessment modalities associated with it) adequately serve countries, and not the other way around. This translates into a double imperative – on the one hand, maximizing the availability and usability of learning data, making sure it is aligned with country needs; on the other hand, complying with the comparability mandate imposed by the SDG indicator

framework. A tension between the principles of in-built comparability and country ownership has thus emerged – the former being better served by large cross-national assessments, the latter more likely to be maximized by national assessments. The negotiation of the reporting protocol for Indicator 4.1.1 has thus put the UIS under considerable strain. A consensus has finally been reached by recourse to a hybrid approach that maximizes data-source flexibility while recognizing the added-value of cross-national assessments.

In Chapter 6, I also argue that one of the most immediate consequences of the impetus given to learning data has been the transformation of the UIS' relative position within the global education field. If UNESCO's statistical efforts were once characterized as a "sad story" (Heyneman, 1999), this chapter in the history of UIS is definitely brought to a close with the advent of SDG4. The UIS has thus gained substantial visibility and centrality in an area in which it had long struggled to affirm its leadership – that is, the production of globally comparable data. This reinvigoration appears to be largely the result of the UIS' ability to posit itself as an honest broker driven by a public-service *ethos* – one able to convene different parties and interests, and to build consensus in a fraught arena riddled with vested interests. Interestingly, the UIS has succeeded in this by bringing to the fore (rather than concealing) the politicized nature of the process, as well as the necessarily imperfect nature of global datasets. While the construction of "good enough data" typically occurs in the backroom (as a means to confer an appearance of objectivity on the data), the UIS has succeeded in bolstering its own credibility by doing exactly the opposite – that is, by exposing messiness, complexity and political difficulties behind the production of globally comparable learning data. At the same time, the chapter alludes to the fact that the UIS' newfound position is by no means a solid one. This is a consequence of the emergence of parallel measurement projects (the World Bank's Human Capital Index being a case in point), but also of the limited institutional alignment between the UIS and UNESCO.

## 7.2. Contributions to the literature

The empirical insight gained from this thesis touches on a range of wider themes relating to the practice of norm-setting and the production of global data. The following sections discuss the broad implications of the dissertation's findings in relation to these two themes.

### 7.2.1. Global goal-setting as an instance of normative tension: a contribution to the norm-dynamics literature

IR scholars have drawn attention to the potential of global goals to trigger and culminate normative shifts. In the light of such observations, this dissertation aims to gain insight into the role played by IOs in processes of normative change – specifically, in relation to

the process of norm emergence and consolidation (Fukuda-Parr and Hulme, 2011). The tortuous journey of the learning agenda has been used as a case study to get an understanding of such dynamics. Thus, the dissertation tracks the salience of learning within the transnational sphere, in order to understand how such processes intersect with broader processes of global goal-setting. The results suggest that, while global goal-setting exercises operate as a key venue for the advance of norms, such processes are far from straightforward. Indeed, in academic literature, the conceptualization of global goals as norm-building exercises has tended to ignore (or gloss over) the possibility for global goals to operate as sites of normative *struggle* and friction. However, the interplay between the SDG4 negotiations and the efforts to advance a learning agenda suggest this might well be the case.

As discussed in Chapter 5, certain organizations seized the making of the post-2015 education agenda as an opportunity to bring new ideas into the education-for-development field, and encourage a shift in focus from schooling (or access) to learning. Such a learning agenda was neatly normative in nature since it aimed to redefine the priorities guiding international bureaucracies and countries, and the conceptualization of education quality. To some extent, the promoters of the learning agenda succeeded in their intent – the crafting of SDG4 contributed crucially to precipitating the position of learning outcomes as a priority area in the education-for-development agenda. However, the making of SDG4 also became a stage for confrontation in that it contributed decisively to the articulation of an alternative normative community – one that rallied around an open rejection of the conceptualization of learning outcomes as a proxy for education quality, and engaged in advocacy effort to prevent a potential narrowing in the ambition of the post-2015 agenda. This alternative community conceded that change in development priorities was necessary, but proposed an alternative normative course – one revolving around a broad understanding of quality rather than on learning outcomes. While the promoters of the learning agenda exerted a considerable influence in the early stages of the debate and succeeded in terms of issue-framing, during the process of goal-formulation there were a number of nods to the demands of those favoring a more holistic definition of education quality. Overall, these dynamics are indicative of the fact that, while global goals might operate as norm vehicles and as privileged opportunities for norm emergence, they can also operate as sites of normative struggle, or even contribute to *slow down* the process of normative change initiated elsewhere.

In the light of this, the results of my research might contribute to an emerging strand of IR scholarship that has taken issue with the neglect of episodes of normative contestation within the constructivist literature (e.g., Deitelhoff and Zimmerman, 2020; Krook and True, 2012; Payne, 2001; or Weiner, 2014). This is true of the line of inquiry spearheaded by Bloomfield (2016) and Carpenter (2011), who emphasize the theoretical significance of episodes of normative continuity (that is, occasions in which norms did *not* change despite advocacy efforts). Indeed, the latter calls for greater attention to be paid to the broad spectrum of roles intervening in processes of normative change – what he terms the *norm-dynamics role-spectrum,* a continuum of positions between the norm

entreneur and the norm antipreneur ideal types (Bloomfield, 2016). By analyzing the advocacy efforts oriented toward preventing the advance of the learning agenda, this dissertation adds to the understanding of these dynamics. More specifically, it contributes to gain insight into the impact of *exogenous* factors in shaping episodes of norm contestation.

Thus, most of the empirical accounts grappling with episodes of norm antipreneurship tend to explain the fate of "would-be norms" as a result of the organizational attributes of norm entrepreneurs and norm resisters (i.e., asymmetries in the organizational advantages enjoyed by norm entrepreneurs and norm antipreneurs), as well as in relation to the strategies and tactics deployed by such actors (see in particular the compilation of case studies edited by Bloomfield and Scott, 2016). However, more limited attention has been paid to the institutional venues and the broader normative environment in which episodes of norm antipreneurship or norm rivalry take place. The importance of extrinsic factors has certainly been highlighted by some of the most prominent theorizers of norm dynamics (see Krook and True, 2012; Müller, Fey and Rauch, 2013; or Wunderlich, 2013, for a discussion of the relevance of exogenous factors and windows of opportunity as key factors shaping norm emergence). Nevertheless, such considerations do feature less prominently in empirical descriptions of episodes of normative contestation and norm bargaining. In other words, it is less clear how external factors affect the interaction between norm promoters and norm resisters, and the final outcome resulting from normative struggles.

The study of the learning turn in the SDG4 context provides empirical insight into the importance of the institutional sites in which normative struggles crystallized, as well as of the effects of the broader normative environment. In relation to the former, the results of the dissertation suggest that formalized and participatory environments articulated as multi-issue *fora* are particularly favorable to strategic compromises and mutual concessions between constituencies with diverging normative preferences. This is the result of the fact that such institutional features tend to blur and weaken power asymmetries between norm entrepreneurs and norm antipreneurs, while creating room for bargaining and negotiation. These are the dynamics observed in the context of the deliberations that took place in the context of the EFA consultations and the WEF 2015 – which placed *on an equal footing* organizations that, in fact, enjoyed unequal degrees of ideational power and advocacy capacity. Thus, the open and inclusive nature of the post-2015 process rendered it possible to (temporarily) overcome the power asymmetries between those favoring a learning turn and those opposing it. Therefore, the *compromise solution* reached during the negotiation of SDG4 cannot be understood as an accurate reflection of the ideational strength and advocacy power of these two normative communities – but does reflect, and is mediated by, the effects of the institutional venues in which the struggle took place.

In relation to the broader normative environment, the SDG4 case study also indicates that the holistic and aspirational approach favored by the Sustainable Development paradigm

created a powerful opportunity for those constituencies opposed to the prospect of a narrowly-framed learning goal. Thus, the SDGs were explicitly oriented towards avoiding the pitfalls of the MDGs (including the prioritization of a number of issues *at the expense of everything else*) – and this narrative proved a useful counter-frame to CSOs and NGOs concerned with the possibility of an education goal single-focused on the improvement of learning outcomes. Thus, the idea of a quality agenda attentive to the multiple dimensions of education quality gained a foothold partially because it fit with the prevailing values within development circles. Once again, the success of the normative labor performed by certain IOs cannot be solely attributed to their persuasive ability or other organizational attributes. Overall, the results of the dissertation suggest that norm-dynamics owe much to exogenous or environmental factors. Consequently, the fate of global norms cannot be approached as a proxy or mechanical reflection of the influence and authority of norm resisters and norm promoters.

The results of the dissertation also add to our understanding of the motivation and the driving forces behind episodes of norm contestation. Norm contestation is conventionally portrayed as the result of a fundamental disagreement regarding the content or substance of norms, i.e., the paradigm or ideational framework they represent, or the anticipated policy implications. However, the debate around the question of learning outcomes suggests that, in some cases, the very organizational identity of norm entrepreneurs is at the root of normative contestation and resistance. My interviews with CSO organizers indicate that the animosity around the prospect of placing learning outcomes at the center of the new development framework owes much to the fact that this move was perceived as aligned with the World Bank's programmatic priorities, or as part of a donor agenda overrepresenting the interests of the Global North. Since some of these CSO and NGOs perceived that their organizational mandate was to keep "instrumentalist" approaches to education at bay, they tended to adopt a vigilant, combative attitude towards proposals advanced by those organizations considered to epitomize this instrumentalist approach (namely, the World Bank, but also certain donors driven by an outcome-oriented philosophy). Hence, the ultimate source of concern around the learning turn did not lie so much in the principles guiding the agenda, but in the profile of its promoters. The debate around learning thus reflected broader concerns about the structure of power of the global education field. In other words, disputes around the learning turn functioned as a *surrogate* debate mirroring tensions on the balance of power between the EFA community and other strands of the post-2015 debate, and on the moral authority and leadership capacity of bilateral and multilateral donors. Such subtleties might explain why concerns around the learning agenda could not easily be transferred to the UN negotiation chambers. It also explains why potentially divisive issues (e.g., the risk of a learning agenda devolving into a "testing craze", or to incentivize the establishment of performance-based accountability schemes) were magnified on the part of certain CSO negotiators – sometimes, very much to the surprise of some promoters of the learning turn who themselves were opposed to such policies.

In this sense, the results of this research suggest that the study of international norm-building can greatly benefit from a relationalist perspective (e.g., field theory). Empirical observations indicate that normative struggles cannot be explained exclusively by reference to the essence of norms, but need to be understood in relation to the social space where such contestation takes place. Thus, the normative preferences of different actors appear to be informed and shaped by their own position within a given space, and reflect their appetite to preserve (or modify) the balance of power. The results of the research are indicative of the fact that IOs do not espouse norms in a vacuum – institutional legacies, including inter-organizational misgivings and attempts at organizational differentiation, are a key driving force within IOs' processes of normative configuration. Ultimately, the position of different IOs in relation to emerging norms cannot be explained solely by the essence of such norms, the (anticipated) impact of normative shifts within a given field of forces must also be included.

### 7.2.2. Ambiguities on the ultimate use(r)s of global datasets: a contribution to the practice of international statistics

The advent of the SDGs has spurred new interest in the impact and making of international statistics, the political economy of statistical capacity, and the global governance of statistics. This is an important move in that, as noted in Chapter 3, literature on global indicators has tended to privilege comparative and evaluative formats (i.e., ranking and rating practices), with limited attention being paid to more mundane or routine forms of quantification – including the development of the international statistical system, or the intersection between development efforts and statistical capacity-building (see Ward, 2004, for an exception). Remarkably, with respect to SDGs, debates on the challenges posed by quantification have not been confined to academic circles – as had been largely the case for the discussions around the MDG measurement efforts. Thus, the advent of the Data Revolution associated with the post-2015 agenda have also generated important debates within practitioner circles. A number of development partners have started to reflect on the costs, risks and trade-offs involved in the growing demand for development data. While it is widely recognized that increased data availability is a powerful enabler of progress, there is also a growing awareness that, on the one hand, the potential of data is only realized under certain circumstances, and on the other hand, not all forms of development data are equally likely to serve the objectives laid out by the 2030 Agenda for Sustainable Development. This line of reasoning is captured by a 2015 report prepared by the Partnership in Statistics for Development in the 21st Century (PARIS21), established by the UN, the European Commission, the OECD, the IMF and the World Bank[176]:

---

[176] In general, the preoccupation over data governance within development circles is also reflected in the establishment of a number of projects (or working groups) concerned with such questions. Some prominent examples of this trend are the project *Governing Data for Development* launched by the Center for Global Development, networks such as the Global Partnership for Sustainable Development Data and the aforementioned PARIS21 initiative, the work of the (self-professed)

It is clear that the indicator and data requirements of any new framework risk being a burden for some developing countries. A concerted effort from the international community over the next 15 years will be needed to ensure that SDG monitoring does not impose inordinate costs on developing countries or divert resources from achieving national statistical development strategies. (PARIS21, 2015, p. 20)

There is indeed mounting evidence that the production of global datasets poses important dilemmas in terms of data sources and data-collection procedures. In the area of health, Kavanagh, Katz and Holmes (2020) have recently drawn attention to the fact that most global datasets are frequently of little use to local decision-makers as they lack the granularity and precision necessary to orient domestic policy-making, resource allocation or service delivery. In addition, in relation to the area of health, Jerven (2018) has argued that "there is a trade-off between precision and relevance" (p. 469), and that, since statistical capacity is a limited resource, such disjunctives merit greater attention than they are being afforded within the development arena. Indeed, the same author has written extensively on the limitations of the SDG indicator framework and its limited alignment to country needs. In a 2017 paper oriented towards assessing the economic and opportunity costs of the measurement agenda associated with the SDGs, he concluded:

> Governments need disaggregated, high-frequency data linked to sub-national units of administrative accountability. In contrast, the SDGs emphasize global goals, standards, and comparability. This emphasis on monitoring progress toward an indicator that supports donor goals essentially lowers the fungibility of the statistics […] A good example is poverty headcount data […] While these poverty data create important baselines and put short-term policy planning into long-term perspective, the danger is that donor preferences for global comparable data come at the expense of the reliable and high-frequency data needed at the local level. (Jerven, 2017, p. 14).

A similar argument has been advanced by MacFeely and Barnat, (2017), who have argued that the expansive nature of the SDG monitoring framework could end up diverting statistical resources from the production of nationally relevant indicators. The authors call for the need to design statistical capacity-building programs that do not trump national and regional priorities. Custer and Sethi's (2017) comments on the disconnect between the supply and demand for development, and the risk that SDGs end up producing "data graveyards", point to similar problems. In the same vein, Sandefur and Glassman (2014) observed that pay-for-performance initiatives might create perverse incentives for country reporting practices, and that "The different needs of donors and government present trade-offs between the comparability, size, scope, and frequency of data collection" (p. 4). Overall, there is a proliferation of studies concerned with the opportunity costs of the measurement needs associated with global goals and international comparability.

This dissertation contributes to this line of inquiry by examining one of the measurement areas in which the tensions between global and domestic data needs crystallizing, namely,

---

*watchdog organization* Open Data Watch, or the Thematic Research Network on Data and Statistics launched by the UN Sustainable Development Solutions Network.

the production of learning data. Hence, and as noted in Chapter 5, one of the ideational repercussions of the learning turn has been the coupling of the learning-crisis problem with the learning-measurement solution in the global education agenda. The idea has found major resonance in different IOs that have incorporated the improvement of learning assessment systems as one of their programmatic priorities[177]. Thus, Large-Scale Learning Assessments (LSLAs) appear to have acquired the status of a consensual policy script within the development realm, largely as a result of their potential as a policy tool – that is, a tool useful to spur government action and to inform domestic policy-making (for instance by informing regulatory and resource-allocation activity) (UNESCO, 2019). At the same time, the measurement needs associated with the SDGs indicator framework have given additional impetus to the assessment agenda, as LSLAs constitute necessary data sources for global reporting purposes. However, such duplicity of functions (i.e., the fact that LSLAs operate as a policy tool while serving reporting purposes) is at the root of a number of trade-offs faced by countries, development partners, and the international statistical community. This is so that those assessments more amenable to (or susceptible to be used for) comparative purposes are not necessarily those more appropriate to operate as policy-informing tools.

Thus, national assessments might hold great potential to inform domestic action – they are particularly apposite to monitor learning trends over time, inform the design of policy interventions, or identify priority areas or populations. However, in raw form, they do not necessarily lend themselves to comparative and reporting purposes. Conversely, cross-national assessments are better-suited for global monitoring purposes (in that they are easier to harmonize) and might be uniquely placed to create an impetus for change, establish policy priorities or leverage certain political agendas. Nevertheless, they are unlikely to have the necessary granularity[178] or frequency for policy-planning or policy-design purposes (Lockheed, 2012, 2016; Tobin, Lietz, Nugroho, Vivekanan and Nyamkhuu, 2015; Wagner, Wolf and Boruch, 2018; Winthrop and Anderson, 2013). Thus, the comparability imperative associated with global reporting might end up emptying the LSLAs program from its policy-planning potential – if it incentivizes countries and donors to privilege a specific subset of assessments administered externally.

---

[177] While the growing prominence of LSLAs within the agendas of different IOs is beyond the scope of this dissertation, it should be noted that, since the adoption of the Education 2030 agenda, organizations such as the GPE, the World Bank or the UNESCO-IIEP have all put in place (or strengthened substantially) a series of programs oriented towards improving the availability, quality and actionability of data on learning outcomes.

[178] For instance, cross-national assessments are likely to be affected by so-called *floor effects* (when most of the participants in a given assessment attain scores close to the baseline) or *ceiling effects* (when most of the participants obtain the maximum score). Similarly, in order to be useful for policy or resource-allocation purposes, disaggregation levels need to coincide with units of accountability and management (provinces, municipalities, schools). Cross-national assessments rarely lend themselves to low levels of disaggregation – not least because they are typically sample-based rather than census-based.

As a result of these trends, and as discussed in Chapter 6, the negotiation of the reporting protocol for Indicator 4.1.1. has been shaped by a recurring tension between in-built comparability and country-ownership as two basic principles expected to orient the production of SDG4 data. The UIS and GAML have indeed gone to great lengths to maximize data-source flexibility and ensure that the SDG4 monitoring needs do not distort country efforts to strengthen systems of learning measurement. However, striking such a balance has been far from a straightforward process – technical and political challenges have been manifold and persistent.

These results have implications (or might provide useful insight) for international development efforts. The findings suggest that the unintended effects of global reporting requirements, and the risk of an over-emphasis on global comparability, warrant further consideration on the part of donors or aid programs supporting countries in their measurement efforts. Thus, development partners providing economic or technical assistance for statistical capacity-building have a key role to play in making sure that efforts to populate the global indicators (including the SDGs, but also the Human Capital Index) do not occur at the expense of the data needed to inform and support national policy-making. Similarly, countries need to be informed of the opportunities and costs, pluses and minuses associated with different forms of LSLAs. The responsibility to help countries navigate the "assessment market" (cf. Montoya and Crouch, 2019) cannot fall exclusively on the UIS. Too many donors and development partners continue to address LSLAs as a unitary category – with differences in the potential of cross-national and national assessments being largely glossed over or addressed in a cursory way. In fact, these two varieties are often lumped together into a single category, following the classification popularized by Clarke (2012) in the context of the World Bank's SABER project. Guidance on the specificities of different varieties of LSLAs therefore need to be provided in a more systematic and objective way. This is to ensure that countries make informed decisions aligned with domestic needs rather than driven by the preferences of development partners in charge of technical and financial support, or by symbolic or reputational concerns (e.g., an interest in participating in a "global ritual of belonging", cf. Addey, 2014). Finally, the results of the dissertation suggest that is necessary to (continue to) work on the development of measurement strategies that make possible the use of national, population-based or bottom-up assessments for global reporting purposes. Even if this strategy results in a loss of technical rigor in terms of global comparability, it appears to be the most promising option in order to ensure that resource-constrained countries are not forced to choose between global reporting and the development of autonomous learning assessment systems.

Beyond practical considerations relative to the development of learning assessment systems, the results of the dissertation also contribute to more theoretically-oriented academic debates by shedding light on the process of *production* of global data. The production of data represents the second stage in the indicator trajectory advanced by Davis, Kingsbury and Merry (2015) – however, it remains a comparatively under-researched phase, in contrast with the conceptualization, use and impact of indicators,

which are the object of a burgeoning literature. Particularly in relation to the SDGs indicator framework, the bulk of the research has focused on the formulation of indicators or on reporting procedures. Conversely, the process through which indictors are linked to existing or new data sets has received limited attention – see for instance the special issues on the intersect between SDGs, data and knowledge, recently published by *Global Policy* (Vol. 20, Issue S1 – 2019) and by the *Journal of Education Development and Capabilities* (Vol. 20, Issue 4 – 2019). By focusing on the production stage, the dissertation contributes to an understanding of the collective (and conflict-ridden) nature of indicator-making. It highlights specifically that data suppliers approach the assemblage of global data as an opportunity to consolidate (or even expand) their portfolio of activity and areas of influence. As a result, the orchestration labor performed by IOs in charge of definition reporting and harmonization standards represents a particularly challenging enterprise. The role of the UIS in the production of SDG4 learning data illustrates the difficulties in keeping assessment producers on board (offering them reputational or material incentives powerful enough to get them to share their data) while "taming" the assessment industry – for instance, preventing certain data producers from imposing their methodological preferences, or from using SDG4 as a product placement scheme.

In addition, the dissertation adds to our understanding of the symbolic labor behind the transformation of raw data into authoritative representations – the process through which an indicator "moves upwards" and thus becomes reliable (cf. Espeland and Stevens, 2008). In this sense, it corroborates Rocha de Siqueira's (2017) observations that data producers recognize global datasets as inherently imperfect, but accept such inaccuracies out of pragmatism. However, the dissertation contends that the approximate and error-prone nature of global data is not simply accepted as a "lesser evil" on the part of indicator promulgators – imperfections can even be seen as an opportunity for IOs to reassert their own authority. This is because the production of "good enough" data entails an element of judgement that cannot rely exclusively on technical considerations – a role for which IOs are ideally suited on account of their aura of neutrality and commitment to the common good.

## 7.3. Building consensus in a polycentric environment: Some remarks on the study of multi-arena politics

One of the objectives of this thesis was to gain an understanding of the mechanisms through which IOs with different priorities have reached a consensus regarding the targets and the indicators of the new education agenda. In the process of conducting this research however, this objective proved to be increasingly elusive. This was so for two different reasons. First, many of the agreements reached during the run-up to the adoption of SDG4 can be characterized as bargaining compromises rather than as instances of reasoned consensus. The difference between these two categories is discussed in depth by Thomas Risse's seminal essay *Global Governance and Communicative Action,* which builds on Elster's (1991) and Saretzki's (1996) distinction between arguing and bargaining as two

ideal-type modalities of communication. Risse argues thus that "bargaining compromises refers to cooperative agreements through the give and take of negotiations based on fixed interests and preferences, [*whereas*] reasoned consensus refers to the voluntary agreement about norms and rules reached through arguing and persuasion" (2004, p. 302). Whilst, as noted by Risse, differentiating compromise from consensus is not an easy endeavor and requires careful process tracing, a key difference lies in the fact that, in the case of bargaining, there is no change in the preferences and interests of the different parts.

This seems indeed to be very much the case of the final nature of both the SDG4 targets and the accompanying Framework for Action. As argued above, the final agreement does not appear to hinge on a process of ideational rapprochement and persuasion, but on a series of deals and exchanges, give-and-take, and mutual concessions. Overall, the construction of the agenda appears to revolve around an additive (rather than a transformative) process. This consensus-by-aggregation dynamics is in fact not exclusive to the education sector – the "Christmas tree" nature of the SDGs and the lack of prioritization have been recurrent targets of criticism on the part of public commentators[179].

This additive and compromise-like nature of the SDGs was not, however, the main impediment to an analysis of consensus-building in the context of the post-2015 debate. Even if the notion of consensus was approached in its broadest sense, or even if the thesis turned its focus on compromise-making practices, an additional and more fundamental challenge remained – namely, the fact that SDG4 negotiations often seemed to advance without any kind of deliberation taking place. The notions of arguing or bargaining assume that different actors do effectively engage in an exchange of views and participate in a common forum[180]. However, and as discussed above, the post-2015 debate relied on a multiplicity of *fora*, with different strands of the debate taking place at the same time. This atomization, and the presence of overlapping venues, reduced the opportunities, incentives and need for consensus-making.

Thus, the shift from the MDG/EFA times to the SDG era entailed a transition from a mono-centric to a pluricentric negotiation environment. The negotiation of SDGs was thus a multi-layered process, anchored on a wide variety of venues and arenas. As a result of this polycentric arrangement, different actors and constituencies did not need to engage in argumentative exchanges – in other words, they could afford to push their own agendas in the forum that proved more welcoming of such views, rather than focus on finding common ground. The co-existence of different strands of the debate, each of them giving

---

[179] For instance, in 2015 a piece in the economist portrayed the SDGs as "sprawling and misconceived" and noted that the expansive nature of the agenda "means, in practice, no priorities at all".

[180] It is no coincidence that Risse's observations stem largely from a Habermasian theory of communicative action, which places great emphasis on the importance of public forums as spaces for debate and deliberation.

rise to different proposals and target lists and relying on different time horizons[181], operated as an enabling architecture. This was to prevent divisive issues from crystallizing into conflict (or even to become visible), and it spared participants from having to find a compromise. While the process was punctuated by episodes of conflict or open discussion (e.g., the normative struggle around learning outcomes), these moments were comparatively rare.

Such dynamics rendered it difficult to draw conclusions on the mechanism of consensus-building underpinning the negotiation of the SDG4. In the light of empirical observation, the emphasis on consensus (or compromise) seemed ill-conceived or unhelpful in making sense of the post-2015 process. Conversely, these findings might contribute to, and benefit from, an alternative area of study, relative to the impact of overlapping venues, and the role of forum-shopping in global governance. There is indeed an emerging strand of scholarship dealing with the political consequences of overlapping venues, as well as with the agency exerted by IOs in shaping such *fora* and seeking shoppers (as opposed to more conventional approaches to forum-shopping and forum-shifting practices, which adopt a state-centric perspective) (see for instance Kranke, 2017b; and Murphy and Kellow, 2013). While theorizing about the mechanisms of consensus building is beyond the scope of this dissertation, suffice to say that the conceptual and theoretical framework deployed by emerging literature on the topic represents a promising alternative approach to the study of multi-scalar negotiation processes such as the post-2015 process.

## 7.4. Future research directions

This dissertation has inquired into processes of global target-setting and indicator-making, through a case study of the negotiation of SDG4 and its associated metrics. In the process of conducting this research, a number of new hypothesis and research questions emerged – in light of the empirical observations, but also since scholarship on such questions has only continued to grow in volume and sophistication. This section maps future research directions that follow from the findings of the dissertation, organized along three main themes – namely the political economy of education statistics, SDG4 global coordination mechanisms, and the role of professional networks within the education-for-development field.

---

[181] This fragmentation and multiplicity of (sometimes overlapping venues) can be seen not only in the co-existence of a EFA-led and a UN-led strand of the debate or the proliferation of UN venues, but also in the separation of the SDG debate and the Financing for Development negotiations, the creation of the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs) and its coexistence with the TAG, the E-TAG and the TCG, or the growing evolving rapport between the TCG and the SDG4 Steering Committee, and the GAML and the TCG, *i.a.*

*7.4.1. A political economy of education statistics*

Chapter 5 alluded to the fact that there is a risk of external data demands taking precedence over local and domestic data needs. Such findings raise questions regarding the process through which countries decide on data-collection priorities in a context of limited human and economic resources, and regarding how global forces shape national statistical systems. The development field has placed high expectations upon the potential of the data revolution – yet there is limited understanding of the process through which countries' statistical capacity is developed[182], and how this is impacted by the priorities and action of IOs and donors and by the global statistical architecture. Thus, while data is at the core of current development efforts[183], we have limited insight into how global data demands, along with international statistical capacity-building initiatives, effectively impact national statistical systems.

Specifically in the field of education, there is scant empirical research on the interplay between national statistical offices (or line ministries) and the statistical departments of the major education IOs. We have thus limited evidence on how, and to what extent, the data demands associated with SDG4 have affected data-collection efforts at the national level. Similarly, the impact of global reporting requirements on national stakeholders has rarely been the object of scholarly interrogation. Research on global measurement efforts has tended to privilege international organizations in charge of data harmonization and the conceptualization of indicators. Conversely, the role of the most basic data suppliers (i.e., countries) remains something of a blind spot. My research suggests however, that there is much to be learnt from the study of the interaction between national and international agencies involved in the collection and production of education statistics. An analysis of the cooperation and exchanges between national data experts and dedicated IO bureaucrats can contribute to our understanding of how countries accommodate competing data needs and adopt and/or calibrate data-collection instruments. It can also provide insight into how statisticians grapple with the need to conform to international statistical standards – a relevant question given that research on other areas suggests that this is not, by any means, an easy endeavor (cf. Linisi and Mügge, 2019). Additionally, it might help elucidate how national officers negotiate the credibility of their data and why certain data sources (and countries) come to be trusted, whereas others never do. Finally, this line of inquiry can provide insight into potential

---

[182] An exception to this is the line of work recently initiated by the IADB in relation to the political economy of statistical capacity – see Dargent, Lotta, Mejía and Moncada (2018) or Taylor (2016). However, within academic circles, debates on such questions have been conspicuously lacking.

[183] The growing centrality of such questions can be observed in the different *fora* and mechanisms established by the UN following the adoption of the SDGs (e.g., the UN World Data Forum led by the High-Level Group for Partnership, Coordination and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development), but also in the work of other organizations – for instance, the World Bank's World Development Report 2021 will focus on the role of data in development efforts under the banner "Data for Better Lives".

strategies to prevent global and donor data requirements becoming a burden (or even posing a perverse incentive) for countries.

*7.4.2. Following-up SDG4 global coordination mechanisms*

While the dissertation provides some insight into the practice of goal-setting, it only captures a part of the process – namely, the *negotiation* of the targets and indicator framework. A natural next step is to examine the *implementation* of SDG4 and its impact on the global education architecture. In particular, and as argued in Chapter 5, some interviewees suggested that the Education 2030 agenda had progressively lost its momentum – generating limited interest within most of the major education agencies. Similar concerns have been raised in a recent evaluation of UNESCO's education sector, which noted that "Key external stakeholders consider that the global and regional architecture for supporting implementation of SDG4-Education 2030 is not functioning effectively" (UNESCO Internal Oversight Service, 2019, p. vi). Doubts have also arisen around UNESCO's global leadership, as well as around the SDG4-Education 2030 Steering Committee's coordination capacity. At the same time, the celebration of the 2018 Global Education Meeting was touted as a success and has once again provided proof of the convening power of UNESCO. This pattern suggests that, while SDG4 is of great symbolic value, its potential as a focal point for global education debates is far more uncertain. In this sense, it is not clear to what extent the SDG4 architecture represents a shift away from the dynamics that characterized EFA arrangements. Finally, there is limited discussion of the interplay between the SDG4 global coordination mechanisms and the arrangements on which the UN relies for review and implementation of the SDGs – most notably, the High-Level Political Forum and the United Nations Sustainable Development Group.

Thus, the global SDG4 architecture represents a relevant area of inquiry. More specifically, there are two main themes that might benefit from further research. The first theme is the relative prominence of SDG4 within the institutional agendas, programmatic and funding priorities and of the different SDG4 co-conveners and partners[184]. As noted in Chapter 2, the EFA agenda exerted a rather limited influence on the agendas of the partaking IOs. It might be of interest to compare how the sense of ownership of SDG4 differs across different organizations, as well as to examine how the major education agencies differ in their (re)interpretation and assessment of the new framework, and how (and to what extent) SDG4 is mainstreamed into the agendas of different IOs.

The second theme of interest concerns the interaction of different IOs in the context of the SDG4 global coordination mechanisms – most notably, the SDG4-Education 2030 Steering Committee. The global education architecture has been described as exceedingly

---

[184] Namely, UNESCO, UNICEF, UNDP, UNHCR, UNFPA, UN Women, the World Bank, ILO, OECD and GPE.

decentralized or even "broken" (Burnett, 2019), with a growing number of organizations competing for visibility and authority[185]. It might be of interest to examine how the coordination efforts set up by SDG4 have an impact on such dynamics, and to what extent they have intensified practices of inter-organizational collaboration and exchange. In order to elucidate such questions, a potentially fruitful empirical strategy would be to compare the collaboration mechanisms in the education realm with the coordination mechanisms set up in other development areas[186] – contrasting their intensity and formalization, the bureaucratic routines and forms of exchange on which they rely, and the challenges they face.

### 7.4.3. The role of professional networks within the education-for-development field

On the occasion of the 1990 WCEFA, Colette Chabbott observed that professionalization was one of the key forces at play in the spread of global educational norms, and argued that "the heads of large international development organizations engineered a world education conference but international development professionals played the leading role in determining what would be promoted at that conference" (1998, p. 216). At the same time, the author noted that, unlike the health sector, the education field lacked a professional/intellectual consensus and remained divided over core issues. Over the course of my research, this absence of a professional/intellectual consensus became readily apparent. Not only did IO staffers differ significantly in their normative preferences and the identification of priorities, but they also differed on the very conceptualization of educational progress and the strategies, technologies and resources more likely to deliver it. They also appeared to disagree on fundamental questions such as the modes of knowledge that constitute valid expertise – that is, the sources of epistemic authority. For instance, there were striking differences in the pieces of evidence that different interviewees used to "substantiate" their assertions – the GMR was one of the very few pieces of evidence consistently cited across the board. Conversely, other forms of evidence (like CNAs or econometric analysis) were the object of polarizing views – with some interviewees approaching them as authoritative sources and others disparaging them as unfortunate simplifications. Interviewees also differed in their identification of the "centers of gravity" within the education-for-development field – they expressed different views on the relative authority and centrality of UNESCO, the World Bank and GPE.

---

[185] It should be however noted that such trends are not exclusive to the education field – the international development field appears also to be characterized by a fragmentation of authority and has been described as a loosely coupled regime in which unifying paradigms have limited influence (cf. Babb and Chorev, 2016).

[186] The coordination efforts recently undertaken in the area of health represents, for instance, a potential basis for comparison – given the fact that the global health landscape has also been characterized by high levels of fragmentation as a result of the proliferation of global actors and the limited capacity of WHO to exert an effective leadership (Spicer et al., 2020).

These dynamics were beyond the scope of the dissertation and were not examined in a systematic way. However, my dissertation points to the need to gain a better understanding of the professionals populating the education-for-development field. Thus, there might be great value in engaging with an analysis of the educational background and professional trajectories of those individuals that inhabit these offices, departments and organizations in charge of the negotiation and subsequent implementation of the SDG4 agenda[187]. It would be useful to identify paradigmatic career trajectories and professional typologies, to examine the distribution of these profiles across different organizations and within their respective institutional hierarchy, and to analyze distinct patterns of recruitment currently structuring the education-for-development field.

Overall, this line of inquiry represents an important endeavor given that, while in certain issue-areas there has been a profusion of the impact of professional knowledge and professional cultures (e.g., in relation to the international financial governance[188]), such questions have received very limited attention in the area of education. There is also evidence that transnational professional networks have a determining impact over processes of normative change and the construction of expert authority in transnational arenas (Seabrooke and Henriksen, 2017). A better understanding of professional dynamics within the education field might bring insight into the different epistemic communities that seem to coexist within it. At the same time, they could also add further nuance and complexity to the understanding of the normative struggles that appear to riddle global education efforts.

---

[187] While the contours of this community are difficult to determine, a proxy for this group could be the list of the IO staffers and bilateral representatives that attended the WEF 2015 and the GEM 2018, and/or participated in the post-2015 education consultations.

[188] Prominent examples include Chwieroth (2013), Nelson, (2014), and Seabrooke and Tsingou (2009), and Weaver (2008).

# List of references

Note*: The references marked with an asterisk correspond to those documents gathered for the purposes of documentary analysis.*

---

Aberbach, J. D., & Rockman, B. A. (2002). Conducting and coding elite interviews. *PS: Political Science & Politics*, *35*(4), 673–676.

ACER/UIS. (2017). *Principles of Good Practice in Learning Assessment*. Montreal, Canada: UIS. *

Addey, C. (2014). *Why do countries join international literacy assessments? An Actor-Network Theory analysis with case studies from Lao PDR and Mongolia* (Doctoral Dissertation, University of East Anglia). Retrieved from: https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.633806

Addey, C. (2016). PISA for Development and the sacrifice of policy-relevant data. *Educação & Sociedade, 37*(136), 685–706.

Addey, C. (2018). The assessment culture of International Organisations: "From philosophical doubt to statistical certainty" through the appearance and growth of Internatioinal Large-Scale Assessments. In C. Alarcón, & M. Lawn (Eds.), *Assessment Cultures: Historical Perspectives* (pp. 379–408). Berlin: Peter Lang.

Addey, C., & Sellar, S. (2017). A framework for analysing the multiple rationales for participating in international large-scale assessments. *Compare: A Journal of Comparative and International Education, 47*(3), 434–443.

Addey, C., & Sellar, S. (2019). *Is it worth it? Rationales for (non)participation in international large-scale learning assessments* (Education Research and Foresight Working Papers No. 24). Paris: UNESCO.

Adler-Nissen, R. (2013). Introduction. Bourdieu and International Relations Theory. In R. Adler-Nissen (Ed.), *Bourdieu in International Relations. Rethinking Key Concepts in IR* (pp. 1–23). Abingdon, UK/New York: Routledge.

Adler-Nissen, R. (2015). Relationalism or why diplomats find international relations theory strange. In O. J. Sending, V. Pouliot, & I. B. Neumann, (Eds.), *Diplomacy and the Making of World Politics* (pp. 284–308). Cambridge: Cambridge University Press.

Adler, E., & Pouliot, V. (2011). International practices. *International theory, 3*(1), 1–36.

Ahmed, A. (2014). Squaring the circle: EFA in the post-2015 global agenda. *International Journal of Educational Development, 39*, 64–69.

Alexander, R. (2015). Teaching and learning for all? The quality imperative revisited. *International Journal of Educational Development, 40*, 250–258.

Altinok, N. (2017). *Mind the Gap: Proposal for a Standardised Measure for SDG 4–Education 2030 Agenda* (UIS Information Paper No. 46). Montreal, Canada: UIS. *

Altinok, N., Angrist, N., & Patrinos, H. A. (2018). *Global Data Set on Education Quality (1965–2015)* (Policy Research Working Paper No. 8314). Washington, D.C.: The World Bank. *

Arndt, C. (2008). The politics of governance ratings. *International Public Management Journal, 11*(3), 275–297.

ASPBAE. (2016, May). *Newly constituted SDG-Education 2030 Steering Committee defines it work* (Bulletin). Retrieved from: http://www.aspbae.org/userfiles/july16/2_SDG_Education2030_SC.pdf *

Babb, S. (2003). The IMF in sociological perspective: A tale of organizational slippage. *Studies in Comparative International Development, 38*(2), 3–27.

Babb, S., & Chorev, N. (2016). International organizations: loose and tight coupling in the development regime. *Studies in Comparative International Development, 51*(1), 81–102.

Barnett, M. N., & Coleman, L. (2005). Designing police: Interpol and the study of change in international organizations. *International Studies Quarterly, 49*(4), 593–619.

Barnett, M. N., & Finnemore, M. (1999). The politics, power, and pathologies of international organizations. *International Organization, 53*(4), 699–732.

Barnett, M. N., & Finnemore, M. (2004). *Rules for the World: International Organizations in Global Politics.* Ithaca, NY: Cornell University Press.

Barnett, M. N., & Finnemore, M. (2018). Political approaches. In T. G. Weiss, & S. Daws (Eds.), *The Oxford Handbook of the United Nations* (2nd ed., online resource). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199560103.003.0002

Barnett, M., & Duvall, R. (2005). Power in international politics. *International Organization, 59*(1), 39–75.

Barrett, A. M. (2011a). An education Millennium Development Goal for quality: complexity and democracy. *Compare: A Journal of Comparative and International Education, 41*(1), 145–148.

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

Bartl, W., Papilloud, C., & Terracher-Lipinski, A. (2019). Governing by numbers – Key indicators and the politics of expectations. An introduction. *Historical Social Research/Historische Sozialforschung, 44*(2), 7–43.

Bashir, S. (2009). Catalyzing country ownership and aid effectiveness: role of the education for All-Fast Track Initiative Catalytic Fund. *Prospects*, *39*(2), 147–161.

Beach, D., & Pedersen, R. B. (2013). *Process-Tracing Methods. Foundations and Guidelines*. Ann Arbor, MI: The University of Michigan Press.

Bebbington, A., Guggenheim, S., Olsen, E., & Woolcock., M. (2004). Exploring social capital debates at the World Bank. *Journal of Development Studies 40*(5), 33–64.

Béland, D., & Orenstein, M. A. (2013). International organizations as policy actors: an ideational approach. *Global Social Policy, 13*(2), 125–143.

Bennet, A., & Checkel, J. T. (2015). Process tracing: from philosophical roots to best practices. In A. Bennett, & J. T. Checkel (Eds.), *Process Tracing. From Metaphor to Analytic Tool* (pp. 3–37). Cambridge: Cambridge University Press.

Bergh, G., & Couturier, J. (2013). *A Rough Guide to Emerging Consensus and Divergence in Post-2015 Goal Areas.* London: Overseas Development Institute.

Bermingham, D. (2009). *Scaling up aid for education. Lessons from the Education for All Fast Track Initiative (FTI)* (Background paper prepared for the Education for All Global Monitoring Report 2010). Paris: EFA-GMR/UNESCO.

Bermingham, D. (2011). The politics of global education policy: the formation of the Education for All–Fast Track Initiative (FTI). *Journal of Education Policy, 26*(4), 557–569.

Bernstein, S. (2013). Global environmental norms. In R. Falkner (Ed.), *The Handbook of Global Climate and Environment Policy* (pp. 127–145). West Sussex, UK: Wiley-Blackwell.

Berry, C., Barnett, E., & Hinton, R. (2015). What does learning for all mean for DFID's global education work? *International Journal of Educational Development, 40*, 323329.

Bhuta, N., Malito, D. V., & Umbach, G. (2018). Introduction: Of numbers and narratives – Indicators in global governance and the rise of a reflexive indicator culture. In D. V. Malito, G. Umback, & N. Bhuta (Eds.), *The Palgrave Handbook of Indicators in Global Governance* (pp. 1–29). Cham, Switzerland: Palgrave Macmillan.

Biermann, F., & Kanie, N. (2017). Conclusion: Key Challenges for Global Governance through Goals. In N. Kanie, & F. Biermann (Eds.), *Governing through Goals. Sustainable Development Goals as Governance Innovation* (pp. 295–309). Cambridge, MA/London: The MIT Press.

Biermann, F., Kanie, N., & Kim, R. E. (2017). Global governance by goal-setting: the novel approach of the UN Sustainable Development Goals. *Current Opinion in Environmental Sustainability,* 26–31.

Biermann, R., & Koops, J. A. (2016). Studying relations among international organizations in world politics: Core concepts and challenges. In R. Biermann, & J. A. Koops (Eds.), *The Palgrave Handbook of Inter-Organizational Relations in World Politics* (pp. 1–46). London: Palgrave Macmillan.

Bigo, D. (2005). La mondialisation de l'(in) sécurité? Réflexions sur le champ des professionnels de la gestion des inquiétudes et analytique de la transnationalisation des processus d'(in) sécurisation [*The globalization of (in) security? Reflections on the field of concern management professionals and*

*analysis of the transnationalization of (in) security processes*]. *Cultures & Conflits, 58*, 53–101.

Bigo, D., & Madsen, M. R. (2011). Introduction to symposium "A Different Reading of the International": Pierre Bourdieu and international studies. *International Political Sociology, 5*(3), 219–224.

Birdsall, N., & Vaishnav, M. (2005). Education and the MDGS: realizing the Millennium Compact. *Journal of International Affairs*, 58(2), 257–264.

Birdsall, N., Bruns, B., & Madan, J. (2016). *Learning Data for Better Policy: A Global Agenda* (CGD Policy Paper No. 96). Washington, D.C.: Center for Global Development. *

Bissio, R. (2003). Civil Society and the MDGs. *Development Policy Journal, 3*, 151–160.

Björkdahl, A. (2002). Norms in international relations: some conceptual and methodological reflections. *Cambridge Review of International Affairs, 15*(1), 9–23.

Black, R., & White, H. (2003). Millennium Development Goals. A drop in the ocean? In R. Black, & H. White (Eds.), *Targeting Development: Critical Perspectives on the MDGs*. London: Routledge.

Bleich, E., & Pekkanen, R. (2013). How to report interview data. In L. Mosley (Ed.), *Interview Research in Political Science* (pp. 84–105). Ithaca, NY/London: Cornell University Press.

Bloomfield, A. (2016). Norm antipreneurs and theorising resistance to normative change. *Review of International Studies, 42*(2), 310–333.

Bloomfield, A., & Scott, S. V. (2016). *Norm Antipreneurs and the Politics of Resistance to Global Normative Change.* Abingdon, UK/New York: Routledge.

Bourdieu, P. (1993). *Sociology in Question*. London: SAGE.

Bray, M. (1986). If UPE is the answer, what is the question? A comment on weaknesses in the rationale for universal primary education in less developed countries. *International Journal of Educational Development, 6*(3), 147–158.

Brenner, R. (2015). Global goal-setting: how the current development goal model undermines international development law. *Michigan State International Law Review, 24*(1), 145–210.

Broad, R. (2006). Research, knowledge, and the art of "paradigm maintenance": the World Bank's Development Economics Vice-Presidency (DEC). *Review of International Political Economy, 13*(3), 387–419.

Broome, A., & Quirk, J. (2015). Governing the world at a distance: the practice of global benchmarking. *Review of International Studies, 41*(5), 819-841.

Broome, A., & Quirk, J. (2015a). Governing the world at a distance: the practice of global benchmarking. *Review of International Studies, 41*(5), 819–841.

Broome, A., & Quirk, J. (2015b). The politics of numbers: the normative agendas of global benchmarking. *Review of International Studies, 41*(5), 813–818.

Broome, A., & Seabrooke, L. (2012). Seeing like an international organisation. *New Political Economy, 17*(1), 1–16.

Broome, A., Homolar, A., & Kranke, M. (2018). Bad science: international organizations and the indirect power of global benchmarking. *European Journal of International Relations, 24*(3), 514–539.

Browne, S., & Weiss, T. G. (2014). The future UN development agenda: contrasting visions, contrasting operations. *Third World Quarterly, 35*(7), 1326–1340.

Browne. S. (2017). *Sustainable Development Goals and UN Goal-Setting*. London/New York: Routledge.

Burnett, N. (2012). Will there be education goals after 2015? Education's fragile position on the global agenda. In *Commonwealth Education Partnerships 2012/13* (pp. 25–26). London: Commonwealth Secretariat. *

Burnett, N. (2019). Invited Essay. It's past time to fix the broken international architecture for education. *International Journal of Educational Development, 68*, 15–19.

Burnett, N. & Felsman, C. (2012). *Post-2015 education MDG*. London: Overseas Development Institute.

Büthe, T. (2012). Beyond supply and demand: a political-economic conceptual model. In K. E. Davis, A. Fisher, B. Kingsbury, & S. E. Merry (Eds.), *Governance by Indicators. Global Power through Quantification and Rankings* (pp. 29–51). Oxford: Oxford University Press.

Cambridge Education/Mokoro/Oxford Policy Management. (2010). *Mid-term evaluation of the EFA Fast Track Initiative. Final report volume 1, Report Executive Summary and Volume 2 – Annexes*. Cambridge: FTI Evaluation Team. Retrieved from: https://www.government.nl/documents/reports/2010/02/01/mid-term-evaluation-of-the-efa-fast-track-initiative-volume-1-main-report

Carpenter, R. C. (2011). Vetting the advocacy agenda: network centrality and the paradox of weapons norms. *International Organization*, 65(1), 69–102.

Chabbott, C. (1998). Constructing educational consensus: international development professionals and the World Conference on Education for All. *International Journal of Educational Development, 18*(3), 207–218.

Chabbott, C. (2003). *Constructing Education for Development. International Organizations and Education for All*. London/New York: Routledge Falmer.

Chabbott, C. (2013). *Institutionalizing Health and Education for All. Global Goals, Innovations and Scaling Up*. New York: Teachers College Press.

Checkel, J. T. (1997). International norms and domestic politics: bridging the rationalist-constructivist divide. *European Journal of International Relations, 3*(4), 473–495.

Checkel, J. T. (2015). Mechanisms, process, and the study of international institutions. In A. Bennett, & J. T. Checkel (Eds.), *Process Tracing. From Metaphor to Analytic Tool* (pp. 74–97). Cambridge: Cambridge University Press.

Christensen, R. C. (2017 - forthcoming). *Transnational professional competition in fields and ecologies*. Manuscript - Copenhagen Business School. https://doi.org/10.31235/osf.io/euqr3. Retrieved from: https://osf.io/preprints/socarxiv/euqr3/

Chwieroth, J. (2010). *Capital Ideas: The IMF and the Rise of Financial Liberalization*. Princeton, NJ: Princeton University Press.

Chwieroth, J. M. (2008). Organizational change "from within": exploring the World Bank's early lending practices. *Review of International Political Economy, 15*(4), 481–505.

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

Clarke, M. (2012). What Matters Most for Student Assessment Systems: A Framework Paper (SABER Working Paper Series No. 1). Washington, D.C.: The World Bank. *

Clegg, L. (2015). Benchmarking and blame games: exploring the contestation of the Millennium Development Goals. *Review of International Studies, 41*(5), 947–967.

Clemens, M. A. (2004). *The Long Walk to School: International education goals in historical perspective* (CGD Working Paper No. 37). Washington, D.C.: Center for Global Development.

Cohen, A. (2018). Pierre Bourdieu and International Relations. In T. Medvetz, & J. J. Sallaz (Eds.), *The Oxford Handbook of Pierre Bourdieu* (pp. 200–248). Oxford: Oxford University Press.

Colclough, C. (2005). Rights, goals and targets: how do those for education add up? *Journal of International Development, 17*(1), 101–111.

Collier, D. (2011). Understanding process tracing. *PS: Political Science & Politics, 44*(4), 823–830.

Collier, D., Brady, H. E., & Seawright, J. (2004). Sources of leverage in causal inference: toward an alternative view of methodology. In H. E. Brady, & D. Collier (Eds.), *Rethinking Social Inquiry. Diverse Tools, Shared Standards* (pp. 229–266). Lanham MD: Rowman & Littlefield.

Cooley, A. (2015). The emerging politics of international rankings and ratings. A framework for analysis. In A. Cooley, & J. Snyder (Eds.), *Ranking the World. Grading States as a Tool of Global Governance* (pp. 1–38). Cambridge: Cambridge University Press.

Coppedge, M. (1999). Thickening thin concepts and theories: combining large-N and small in comparative politics. *Comparative Politics, 31*(4), 465–76.

Cornut, J. (2017). The practice turn in International Relations theory. In *Oxford Research Encyclopedia of International Studies* (online resource). Oxford: Oxford University Press. doi: 10.1093/acrefore/9780190846626.013.113

Crouch, L., & J.-M. B. (2017, July 6). *Measure learning now to reduce inequality tomorrow (*Blog post – GPE). Retrieved from: https://www.globalpartnership.org/blog/measure-learning-now-reduce-inequality-tomorrow *

CUE. (2011a). *Frequently Asked Questions on the Global Compact on Learning and Global Education*. Center for Universal Education at Brookings. Retrieved from: https://www.brookings.edu/wp-content/uploads/2016/06/20110615_global_compact_qa.pdf *

CUE. (2011b). *A Global Compact on Learning. Taking Action on Education in Developing Countries.* Washington, D.C.: Center for Universal Education at Brookings. *

Cussó, R. (2006). Restructuring UNESCO's statistical services. The "sad story" of UNESCO's education statistics: 4 years later. *International Journal of Educational Development, 26*(5), 532–544.

Cussó, R. (2019). *Les statistiques internationales et le développement : de l'universalisme à la fragmentation* [International statistics and development: from universalism to fragmentation] (UMR-DR Working Paper No. 3). Paris: Université Paris 1 Panthéon Sorbonne/Institute de Recherche pour le Développement.

Cussó, R., & d'Amico, S. (2005). From development comparatism to globalization comparativism: Towards more normative international education statistics. *Comparative Education, 41*(2), 199–216.

Custer, S., & Sethi, T. (2017). Avoiding data graveyards: insights from data producers and users. In S. Custer, & T. Sethi (Eds.), *Avoiding Data Graveyards: Insights from Data Producers & Users in Three Countries* (pp. 1–8). Williamsburg, VA: AidData at the College of William & Mary.

Dahler-Larsen, P. (2014). Constitutive effects of performance indicators: getting beyond unintended consequences. *Public Management Review, 16*(7), 969–986.

Daniel, J. (2010). Reflections on UNESCO's World Education Reports and the Global Monitoring Reports on Education for All. *Norrag News* No. 43, 42–44.

Dargent, E., Lotta, G., Mejía, J. A., & Moncada, G. (2018). *The Political Economy of Statistical Capacity in Latin America.* Washington, D.C.: Inter-American Development Bank.

Data Revolution Group. (2014). *A World That Counts. Mobilising the Data Revolution for Sustainable Development.* New York: Independent Expert Advisory Group on a Data Revolution for Sustainable Development. *

Davis, K. E., Kingsbury, B., & Merry, S. E. (2012). Introduction: Global governance by indicators. In K. E. Davis, A. Fisher, B. Kingsbury, & S. E. Merry (Eds.),

*Governance by Indicators Global Power through Quantification and Rankings* (pp. 3–28). Oxford: Oxford University Press.

Davis, K. E., Kingsbury, B., & Merry, S. E. (2015). Introduction: The local-global life of indicators: Law, power, and resistance. In S. E. Merry, K. E. Davis, & B. Kingsbury (Eds.), *The Quiet Power of Indicators: Measuring Governance, corruption, and Rule of Law* (pp. 1–24). New York: Cambridge University Press.

Deakin, H., & Wakefield, K. (2014). Skype interviewing: reflections of two PhD researchers. *Qualitative research, 14*(5), 603–616.

Deitelhoff, N., & Zimmermann, L. (2020). Things we lost in the fire: how different types of contestation affect the robustness of international norms. *International Studies Review, 22*(1), 51–76.

Desrosières, A. (1998). *The Politics of Large Numbers. A History of Statistical Reasoning*. Cambridge, MA/London: Harvard University Press.

Desrosières, A. (2001). How real are statistics? Four possible attitudes. *Social Research, 68*(2), 339-355.

Desrosières, A. (2015). Retroaction: how indicators feed back onto quantified actors. In R. Rottenburg, S. E. Merry, S.-J. Park, & J. Mugler (Eds.), *The World of Indicators. The Making of Governmental Knowledge through Quantification* (pp. 329–353). Cambridge: Cambridge University Press.

Deterding, N. M., & Waters, M. C. (2018). Flexible coding of in-depth interviews: a twenty-first-century approach. *Sociological Methods & Resaerch.* Ahead-of-print. doi: 10.1177/0049124118799377

Dexter, L. A. (2006). *Elite and Specialized Interviewing*. Colchester, UK: ECPR Press.

Dezalay, Y., & Garth, B. G. (1996). *Dealing in Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order*. Chicago, IL: University of Chicago Press.

Dezalay, Y., & Garth, B. G. (2002). *The Internationalization of Palace Wars: Lawyers, Economists, and the Contest to Transform Latin American States*. Chicago, IL: University of Chicago Press.

DfID. (2010). *Learning For All. DfID's Education Strategy 2010-2015*. London: DfID. *

DfID. (2018). *Better Education Statistics and global Action to improve learning (BESTA). Annual Review.* DfID. Retrieved from: https://devtracker.fcdo.gov.uk/projects/GB-1-204695/documents *

DfID. (2019). *Better Education Statistics and global Action to improve learning (BESTA). Annual Review.* DfID. Retrieved from: https://devtracker.fcdo.gov.uk/projects/GB-1-204695/documents *

Diaz-Bone, R., & Didier, E. (2016). The sociology of quantification – Perspectives on an emerging field in the social sciences. *Historical Social Research, 41*(2), 7–26.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48,*147–160.

Djelic, M.-L., & Sahlin-Anderson, K. (2006). Introduction: a world of governance: the rise of transnational regulation. In M.-L. Djelic, & K. Sahlin-Andersn (Eds.), *Transnational Governance. Institutional Dynamics of Regulation* (pp. 1–28). Cambridge: Cambdige Univesity Press.

Dodds, F., Donoghue, D., & Roesch, J. L. (2017) *The Sustainable Development Goals. A Transformational Agenda for an Insecure World*. London/New York: Routledge.

Easterly, W. (2015). The trouble with the sustainable development goals. *Current History, 114*(775), 322–324.

ECOSOC. (2017). *Report of the High-level Group for Partnership, Coordination and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development* (E/CN.3/2017/3). New York: ECOSOC. *

Education Commission. (2016). *The Learning Generation. Investing in Education for a Changing World. A Report by the International Commission on Financing Global Education Opportunity.* New York: The International Commission on Financing Global Education Opportunity. *

Edwards Jr, D. B., Okitsu, T., da Costa, R., & Kitamura, Y. (2018). Organizational legitimacy in the global education policy field: learning from UNESCO and the Global Monitoring Report. *Comparative Education Review, 62*(1), 31–63.

EFA-GMR/UNESCO. (2012). *EFA Global Monitoring Report 2012. Youth and Skills. Putting Education to Work*. Paris: EFA-GMR/UNESCO. *

EFA-GMR/UNESCO. (2015). *Education for All 2000-2015: Achievements and Challenges*. UNESCO: Paris: EFA-GMR/UNESCO. *

Emirbayer, M. (1997). Manifesto for a relational sociology. *American Journal of Sociology, 103*(2), 281–317.

Emmerji, L., Jolly, R., & Weiss, T.G. (2001). *Ahead of the Curve? UN Ideas and Global Challenges*. Bloomington, IN: Indiana University Press.

Erkkilä, T., & Piironen, O. (2018). *Rankings and Global Knowledge Governance. Higher Education, Innovation and Competitiveness*. Cham, Switzerland: Palgrave Macmillan.

Espeland, W. N., & Stevens, M. L. (2008). A sociology of quantification. *European Journal of Sociology/Archives Européennes de Sociologie, 49*(3), 401–436.

ESRC. (2015). ESRC *Framework for research ethics. Updated January 2015*. ESRC. Retrieved from: https://esrc.ukri.org/files/funding/guidance-for-applicants/esrc-framework-for-research-ethics-2015/

Farquharson, K. (2005). A different kind of snowball: identifying key policymakers. *International Journal of Social Research Methodology*, *8*(4), 345–353.

Faul, M. V. (2014). Future-perfect/present-imperfect: contemporary global constraints on the implementation of a post-2015 education agenda. *International Journal of Educational Development, 39*, 12–22.

Faul, M. V., & Packer, S. (2015). *The Role of Global EFA Architectures* (Background paper prepared for the Education for All Global Monitoring Report 2015). Paris: EFA-GMR/UNESCO.

Filmer, D., Hasan, A., & Pritchett, L. (2006). *A millennium learning goal: measuring real progress in education* (CGD Working Paper No. 97). Washington, D.C.: Center for Global Development. *

Finnemore, M. (1993). International organizations as teachers of norms: the United Nations Educational, Scientific, and Cutural Organization and science policy. *International Organization, 47*(4), 565–597.

Finnemore, M. (1996). Norms, culture, and world politics: Insights from sociology's institutionalism. *International Organization, 50*(2), 325–347.

Finnemore, M. (2007). The politics of global partnership. In J. M. Brinkerhoff, S. C. Smith, & H. Teegen (Eds.), *NGOs and the Millennium Development Goals. Citizen Action to Reduce Poverty* (pp. 103–111). New York/Basingstoke, UK: Palgrave Macmillan.

Finnemore, M., & Hollis, D. B. (2016). Constructing norms for global cybersecurity. *American Journal of International Law, 110*(3), 425–479.

Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. *International Organization*, *52*(4), 887–917.

Finnemore, M., & Sikkink, K. (2001). Taking stock: the constructivist research program in international relations and comparative politics. *Annual Review of Political Science, 4*(1), 391–416.

Fligstein, N., & McAdam, D. (2012). *A Theory of Fields*. Oxford: Oxford University Press.

Florini, A. (1996). The evolution of international norms. *International Studies Quarterly, 40*(3), 363–389.

Fomerand, J. (1996). UN Conferences: media events or genuine diplomacy? *Global Governance, 2*(3), 361–375.

Fox, O., & Stoett, P. (2016). Citizen participation in the UN sustainable development goals consultation process: Toward global democratic governance? *Global Governance: A Review of Multilateralism and International Organizations, 22*(4), 555–573.

Fredriksen, B. (1981). Progress towards regional targets for universal primary education: a statistical review. *International Journal of Educational Development, 1*(1), 1–16.

Freistein, K. (2016). Effects of indicator use: a comparison of poverty measuring instruments at the World Bank. *Journal of Comparative Policy Analysis: Research and Practice, 18*(4), 366–381.

Fukuda-Parr, S. (2011). Theory and policy in international development: human development and capability approach and the millennium development goals. *International Studies Review*, *13*(1), 122–132.

Fukuda-Parr, S. (2014). Global goals as a policy tool: intended and unintended consequences. *Journal of Human Development and Capabilities, 15*(2–3), 118–131.

Fukuda-Parr, S., & Hulme, D. (2011). International norm dynamics and the end of poverty: understanding the millennium development goals. *Global Governance*, *17*, 17–36.

Fukuda-Parr, S., & Yamin, A. E. (2013). The power of numbers: a critical review of MDG targets for human development and human rights. *Development, 56*(1), 58–65.

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.

Gallopin, G. C. (1996). Environmental and sustainability indicators and the concept of situational indicators. A systems approach. *Environmental Modeling & Assessment, 1*(3), 101–117.

Gaventa, J., & Mayo, M. (2009). *Spanning Citizenship Spaces Through Transnational Coalitions: The Case of the Global Campaign for Education* (Institute of Development Studies at the University of Sussex Working Paper No. 43). Brighton, UK: IDS-University of Sussex.

GCE. (2018). *Global Report 2018*. Johannesburg: Global Campaign for Education. *

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York: Basic Books.

Gellers, J. C. (2016). Crowdsourcing global governance: sustainable development goals, civil society, and the pursuit of democratic legitimacy. *International Environmental Agreements: Politics, Law and Economics, 16*(3), 415–432.

George, A. L., & Bennet, A. (2005). *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Gerring, J. (2006). Single-outcome studies: a methodological primer. *International sociology, 21*(5), 707–734.

Glewwe, P. (2002). Schools and skills in developing countries: education policies and socioeconomic outcomes. *Journal of Economic Literature 40*(2), 436–82.

GPE. (2018). *GPE 2020. Improving Learning and Equity through Stronger Education Systems. Strategic Plan 2016-2020*. Washington, D.C.: Global Partnership for Education. *

Granmo, A., & Fourie, P. (2021). *Health Norms and the Governance of Global Development: The Invention of Global Health.* London/New York: Routledge.

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, 28*(5), 696–709.

Grek, S. (2020). Interdependency in transnational education governance. In G. Fan, & T. S. Popkewitz (Eds.), *Handbook of Education Policy Studies. Values, Governance, Globalization and Methodology* (pp. 309–328).

Gudmundsson, H. (2003). The policy use of environmental indicators-learning from evaluation research. *Journal of Transdisciplinary Environmental Studies, 2*(2), 1–12.

Guzzini, S. (2013). Power. In R. Adler-Nissen (Ed.), *Bourdieu in International Relations. Rethinking Key Concepts in IR* (pp. 78–92). Abingdon, UK/New York: Routledge.

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

Haas, P. M. (2002). UN conferences and constructivist governance of the environment. *Global Governance: A Review of Multilateralism and International Organizations, 8*(1), 73–91.

Hall, N., & Woods, N. (2018). Theorizing the role of executive heads in international organizations. *European Journal of International Relations, 24*(4), 865–886.

Halliday, T. C., & Carruthers, B. G. (2007). The recursivity of law: global norm making and national lawmaking in the globalization of corporate insolvency regimes. *American Journal of Sociology, 112*(4), 1135–1202.

Halliday, T. C., & Shaffer, G. (Eds.). (2015). *Transnational Legal Orders.* Cambridge: Cambridge University Press.

Halliday, T. C., Block-Lieb, S., & Carruthers, B. G. (2009). Rhetorical legitimation: global scripts as strategic devices of international organizations. *Socio-Economic Review, 8*(1), 77–112.

Hanushek, E. & Kimko, D. (2000). Schooling, labor force quality and the growth of nations. *American Economic Review 95*(5), 1184–208.

Hanushek, E., & Woessmann, L. (2008) The role of cognitive skills in economic development. *Journal of Economic Literature 46*(3), 607–68.

Harvey, W. S. (2011). Strategies for conducting elite interviews. *Qualitative research*, *11*(4), 431–441.

Hay, C. (2016). Process tracing: a laudable aim or a high-tariff methodology? *New political economy*, *21*(5), 500–504.

Hewlett Foundation. (2008). *The Quality Education in Developing Countries Initiative Grantmaking Strategy*. Hewlett Foundation. Retrieved from: https://hewlett.org/wp-content/uploads/2016/08/QEDCGrantmakingStrategy_Final.pdf *

Heyneman, S. P. (1999). The sad story of UNESCO's education statistics. *International Journal of Educational Development, 19*, 65–74.

Heyneman, S. P. (2003). The history and problems in the making of education policy at the World Bank 1960–2000. *International Journal of Educational Development, 23*(3), 315–337.

Heyneman, S. P. (2009). The failure of education for all as political strategy. *Prospects, 39*(1), 5–10.

Hoffmann, M. J. (2010). Norms and social constructivism in international relations. In R. A. Danemark, & R. Marlin-Bennett (Eds.), *The International Studies Encyclopedia* (pp. 5410–5426). Hoboken, NJ: Wiley-Blackwell.

Hopf, T. (2010). The logic of habit in international relations. *European Journal of International Relations, 16*(4), 539–561.

Hüfner, K. (2017). The financial crisis of UNESCO after 2011: political reactions and organizational Consequences. *Global Policy, 8*, 96–101.

Hulme, D. (2007). *The Making of the Millennium Development Goals: Human Development Meets Results- based Management in an Imperfect World* (Brooks World Poverty Institute Working Paper No. 16). Manchester, UK: University of Manchester – Brooks World Poverty Institute.

Hulme, D. (2010). *The Millennium Development Goals (MDGs): A Short History of the World's Biggest Promise* (Brooks World Poverty Institute Working Paper No. 100). Manchester, UK: University of Manchester – Brooks World Poverty Institute.

Hulme, D., & Fukuda-Parr, S. (2009). *International Norm Dynamics and "the End of Poverty": Understanding the Millennium Development Goals (MDGs)* (Brooks World Poverty Institute Working Paper No. 96). Manchester, UK: University of Manchester – Brooks World Poverty Institute.

Hulme, D., & Scott, J. (2010). The political economy of the MDGs: retrospect and prospect for the world's biggest promise. *New Political Economy, 15*(2), 293–306.

IAC-WCEFA. (1990). *Final Report. World Conference on Education for All: Meeting Basic Learning Needs*. New York: UNICEF. *

IAEG. (2015). *Summary of comments (25 September 2015) and List of indicator proposals (11 August 2015).* Retrieved from: https://unstats.un.org/sdgs/files/meetings/iaeg-sdgs-meeting-02/Summary%20of%20Comments%20on%20Indicator%20Proposals%20with%20edits%2019-10-15.pdf *

IBE. (2020). *Recommendations 1 to 48*. Geneva: International Bureau of Educaion. Retrieved from: http://www.ibe.unesco.org/fileadmin/user_upload/archive/policy/ICE_PDFs/R01.pdf

IEA. (2018). *IEA's Rosetta Stone: Measuring global progress toward the SDG for quality education by linking regional assessment results to TIMSS and PIRLS International Benchmarks of Achievement* (GAML5/REF/4.1.1-24). Montreal, Canada: UIS. *

IEG-WB. (2006). *From Schooling Access to Learning Outcomes. An Unfinished Agenda. An Evaluation of World Bank Support to Primary Education*. Washington, D.C.: The World Bank.

Ipsos Mori. (2018). *Evaluation of the Global Education Monitoring Report. Final Report v3*. London: Ipsos Mori. Retrieved from: https://www.ipsos.com/sites/default/files/ct/publication/documents/2019-01/evaluation-of-global-education-monitoring-report-october-2018.pdf

Ito, H. (2012). Jomtien to Jomtien: the Evolving Coordination Process of Education for All 1990-2011. *International Education Studies, 5*(5), 1–13.

Jackson, P. (2008). Pierre Bourdieu, the "cultural turn" and the practice of international history. *Review of International Studies, 34*(1), 155–181.

Jackson, P. T., & Nexon, D. H. (1999). Relations before states: substance, process and the study of world politics. *European Journal of International Relations, 5*(3), 29–332.

Jerven, M. (2013). *Poor Numbers. How We Are Misled by African Development Statistics and What to Do about It?* Ithaca/London: Cornell University Press.

Jerven, M. (2017). How much will a data revolution in development cost? *Forum for Development Studies, 44*(1) 31–50.

Jerven, M. (2018). Beyond precision: embracing the politics of global health numbers. *The Lancet/Perspectives, 392*, 468–469.

Jolly, R, Emmerji, L., Ghai, D., & Lapéye, F. (2004). *UN Contributions to Development Thinking and Practice.* Bloomington, IN: Indiana University Press.

Jolly, R. (2004). Global Development Goals: the United Nations experience. *Journal of Human Development, 5*(1), 69–95.

Jolly, R., Emmerji, L., & Weiss, T. G. (2009). *UN Ideas That Changed the World.* Bloomington, IN: Indiana University Press.

Jones, P. W. (1992). *World Bank Financing of Education: Lending, Learning and Development.* London/New York: Routledge.

Jones, P. W. (2007). WCEFA: A moment in the history of multilateral education. *International Perspectives on Education and Society, 8*, 521–538.

Jones, P. W., with Coleman, D. (2005). *The United nations and Education. Multilateralism, Development and Globalisation.* London/New York: Routledge Falmer.

Kamau, M., Chasek, P., & O'Connor, D. (2018). *Transforming Multilateral Diplomacy. The Inside Story of the Sustainable Development Goals*. Abingdon, UK/New York: Routledge.

Katzenstein, P. J. (1996). Introduction: alternative perspectives on national security. In P. Katzenstein (Ed.), *The Culture of National Security. Norms and Identity in World Politics* (pp. 1–32). New York: Columbia University Press.

Kavanagh, M. M., Katz, I. T., & Holmes, C. B. (2020). Reckoning with mortality: global health, HIV, and the politics of data. *The Lancet, 396*, 288–290.

Keck, M., & Sikkink, K. (1998) *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca, NY: Cornell University Press.

Kelley, J. G., & Simmons, B. A. (2014, August-September). *The power of performance indicators: rankings, ratings and reactivity in International Relations*. Paper presented at the Annual Meeting on the American Political Science Association, Washington D.C. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2451319

Kelley, J. G., & Simmons, B. A. (2019). Introduction: the power of global performance indicators. *International Organization, 73*(3), 491–510.

Kentikelenis, A. E., & Seabrooke, L. (2017). The politics of world polity: script-writing in international organizations. *American Sociological Review, 82*(5), 1065–1092.

Kille, K. J. (2013). Secretaries-General of international organizations: leadership capacity and qualities. In B. Reinalda (Ed.), *Routledge Handbook of International Organization* (pp. 218–230). Abingdon, UK/New York: Routledge.

King, K. (2007). Multilateral agencies in the construction of the global agenda on education. *Comparative Education, 43*(3), 377–391.

King, K. (2010). (Ed.). *A world of reports? A critical review of global development reports with an angle son education and training* (NORRAG News No. 43). Geneva: Graduate Institute of International and Development Studies.

King, K. (2016). The global targeting of education and skill: policy history and comparative perspectives. *Compare: A Journal of Comparative and International Education, 46*(6), 952–975.

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.

King, K., & McGrath, S. (2012, June). *Education and development in Africa: Lessons of the past 50 years for beyond 2015*. Paper presented at the CAS@ 50 Conference, Edinburgh, UK. Retrieved from: https://nottingham-repository.worktribe.com/output/710521/education-and-development-in-africa-lessons-of-the-past-50-years-for-beyond-2015

King, K., & Palmer, R. (2013*). Education and skills Post-2015: What Evidence, Whose Perspectives?* (NORRAG Working Paper No. 6). Geneva: NORRAG.

King, K., & Rose, P. (2005). International development targets and education: towards a new international compact or a new conditionality? *Journal of International Development, 17*, 97.

Klotz, A. (1995) *Norms in International Relations: The Struggles against Apartheid*. Ithaca, NY: Cornell University Press.

Kranke, M. (2017a). *The politics of collaborative global governance: organisational positioning in IMF-World Bank collaboration* (Doctoral thesis, University of Warwick). Retrieved from: http://wrap.warwick.ac.uk/101269/

Kranke, M. (2017b). *From Forum Shopping to Forum Shaping. An Overlooked Form of International Organisation Power* (CSGR Working Paper No. 282/17). Coventry, UK: The University of Warwick.

Krasner, S. D. (1982). Structural causes and regime consequences: regimes as intervening variables. *International organization, 36*(2), 185–205.

Krook, M. L., & True, J. (2010). Rethinking the life cycles of international norms: the United Nations and the Global Promotion of Gender Equality. *European Journal of International Relations 18*(1), 103–127.

Kustermans, J. (2016). Parsing the practice turn: practice, practical knowledge, practices. *Millennium, 44*(2), 175–196.

Lavenex, S. (2015). Institutional interplay with international organizations in the external projection of EU rules. In E. da Conceição-Heldt, M. Koch, & A. Liese (Eds.), *Internationale Organisationen Autonomie, Politisierung, interorganisationale Beziehungen und Wandel* (pp. 349–369). Baden-Baden, Germany: Nomos Verlagsgesellschaft.

Leander, A. (2010). Habitus and field. In E. Denemark (Ed.), *International Studies Compendium Project* (pp. 3255–3270). Oxford: Wiley-Blackwell

Lehtonen, M. (2017). Operationalizing information: measures and indicators in policy formulation. In Howlett, & I. Mukherjee (Eds.), *Handbook of Policy Formulation* (pp. 161–181). Cheltenham, UK/Northampton, MA: Edward Elgar.

Lehtonen, M., Sébastien, L., & Bauler, T. (2016). The multiple roles of sustainability indicators in informational governance: between intended use and unanticipated influence. *Current Opinion in Environmental Sustainability, 18*, 1–9.

Leiteritz, R. J., & Moschella, M. (2010). The IMF and capital account liberalization: A case of failed norm institutionalization. In S. Park, & A. Vetterlein (Eds.), *Owning Development. Creating Policy Norms in the IMF and the World Bank* (pp. 163-180). Cambridge/New York: Cambridge University Press.

Lewin, K. M. (2007). *Long Term Planning for EFA and the MDGs: Modes and Mechanisms* (Consortium for Research on Educational Access, Transitions and Equity Research Monograph No. 7). Centre for Universal Education – University of Sussex. Brighton, UK: CREATE.

Lewin, K.M. (2008). *Strategies for Sustainable Financing of Secondary Education in Africa*. Washington D.C.: The World Bank.

Lingard, B., Sam S., & Baroutsis, A. (2015). "¡Researching the habitus of global policy actors in education. *Cambridge Journal of Education*, 45(1), 25–42.

Linsi, L., & Mügge, D. K. (2019). Globalization and the growing defects of international economic statistics. *Review of International Political Economy, 26*(3), 361–383.

Lipson, M. (2007). Peacekeeping: organized hypocrisy? *European Journal of International Relations, 13*(1), 5–34.

Lipson, M. (2016). Organization theory and cooperation and conflict among international organizations. In R. Biermann, & J. A. Koops (Eds.), *The Palgrave Handbook of Inter-Organizational Relations in World Politics* (pp. 67–96). London: Palgrave Macmillan.

Little, A., & Miller, E. (2000). *The International Consultative Forum on Education for All, 1990-2000. An Evaluation. A Report to the Forum's Steering Committee.* Paris: EFA-Forum Secretariat/UNESCO.

Lockheed, M. (2012). Policies, performance and panaceas: the role of international large-scale assessments in developing countries. *Compare: A Journal of Comparative and International Education, 42*(3), 512–518.

Lockheed, M. (2016). *Measures that Matter: Learning Outcome Targets for Sustainable Development Goal 4 - An examination of national, regional and international learning assessments* (Background paper prepared for the 2016 Global Education Monitoring Report - ED/GEMR/MRT/2016/P1/31). Paris: GEMR/UNESCO.

Lockheed, M. E., & Wagemaker, H. (2013). International large-scale assessments: thermometers, whips or useful policy tools? *Research in Comparative and International Education, 8*(3), 296–306.

MacFeely, S., & Barnat, N. (2017). Statistical capacity building for sustainable development: developing the fundamental pillars necessary for modern national statistical systems. *J ournal of the International Association for Official Statistics, 33*(4), 895–909.

Madsen, M. R. (2006). Transnational fields: elements of a reflexive sociology of the internationalisation of law. *RetfÊrd. Nordisk Juridisk Tidsskrift, 29*(3/114), 34–41.

Manning, R. (2009): *Using indicators to encourage development. Lessons from the Millennium Development Goals*. Copenhagen: Danish Institute for International Studies. Retrieved from: https://www.ecbproject.org/help-library/using-indicators-to-encourage-development-lessons-from-the-millennium-development-goals

March, J. G., & Olsen, J. P. (1989) *Rediscovering Institutions: The Organisational Basis of Politics*. New York: Free Press.

March, J. G., & Olsen, J. P. 1998. The institutional dynamics of international political orders. *International Organization 52*(4), 943–969.

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). London: Routledge.

Martin-Mazé, M. (2017). Returning struggles to the practice turn: how were Bourdieu and Boltanski lost in (some) translations and what to do about it? *International Political Sociology, 11*(2), 203–220.

Martin, C. J. (2013). Crafting interviews to capture cause and effect. In L. Mosley (Ed.), *Interview Research in Political Science* (pp. 109–124). Ithaca, NY/London: Cornell University Press.

Martin, J. L. (2003). What is field theory? *American Journal of Sociology, 109*(1), 1–49.

Mason, J. (2002). *Qualitative Researching*. London: SAGE.

McArthur, J. W. (2014). The origins of the Millennium Development Goals. *The SAIS Review of International Affairs, 34*(2), 5–24.

McCourt, D. M. (2016). Practice theory and relationalism as the new constructivism. *International Studies Quarterly, 60*(3), 475–485.

McGrath, S. (2018). *Education and Development*. Abingdon, UK/New York: Routledge.

McNeely, C. L. (1995). Prescribing national education policies: the role of international organizations. *Comparative Education Review, 39*(4), 483–507.

Menashy, F. (2018). Multi-stakeholder aid to education: power in the context of partnership. *Globalisation, Societies and Education, 16*(1), 13–26.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative Research. A Guide to Design and Implementation* (4th edition). San Francisco, CA: Jossey-Bass, Wiley.

Merry, S. E. (2011). Measuring the world indicators. Human rights, and global governance. *Current Anthropology, 52*(3), 583–594).

Merry, S. E. (2016). *The Seductions of Quantification. Measuring Human Rights, Gender Violence, and Sex Trafficking*. London/Chicago IL: The University of Chicago Press.

Merry, S. E. (2017). Expertise and Quantification in Global Institutions. In R. Niezen, & M. Sapignoli (Eds.), Palaes of Hope: The Anthropology of Global Organizations (pp. 152–171). Cambridge: Cambridge University Press.

Merry, S. E. (2019). The Sustainable Development Goals confront the infrastructure of measurement. *Global Policy, 10*, 146–148.

Momani, B. (2005) Limits of streamlining fund conditionality: IMF's organizational culture. *Journal of International Relations and Development, 8*(2), 142–63.

Montoya, S. (2015). *Monitoring Education for All Global and Thematic Indicators. ETAG Report to the EFA Drafting Group.* Presentation delivered by the UIS Director. *

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 *

Montoya, S. (2018a). *Upgrade SDG indicator 4.1.1(a).* Presentation delivered in the 8th Meeting of the IAEG-SDG (November 2018, Stockholm). Retrieved from: https://unstats.un.org/sdgs/files/meetings/iaeg-sdgs-meeting-08/4.5%20UNESCO-UIS%204.1.1a%20Reclassification.pdf *

Montoya, S. (2018b, October 16). *It is time to upgrade the SDG 4 learning indicator for young children* (Blog post – World Education Blog). Retrieved from: https://gemreportunesco.wordpress.com/2018/10/16/it-is-time-to-upgrade-the-sdg-4-learning-indicator-for-young-children/ *

Montoya, S. (2018bc, November 9). *We are ready to start monitoring early-grade learning* (Blog post – UIS). Retrieved from: http://uisunesco.org/fr/node/1658956 *

Montoya, S. & Hastedt, D. (2017, September 28). *News from Hamburg: big steps forward towards reliable metrics to harmonise learning assessment data* (Blog post – NORRAG). Retrieved from: https://www.norrag.org/news-hamburg-big-steps-forward-towards-reliable-metrics-harmonise-learning-assessment-data-globally-silvia-montoya-dirk-hastedt/ *

Montoya, S., & Crouch, L. (2019, April 26 and May, 20). *The learning assessment market: pointers for countries* (Part 1 and 2). (Blog post – GEMR). Retrieved from: https://gemreportunesco.wordpress.com/2019/04/26/the-learning-assessment-market-pointers-for-countries-part-1/ *

Montoya, S., & Naidoo, J. (2019, August 26). *Learning counts: we must keep children's learning on the 2030 Agenda* (Blog post – UIS). Available at: https://sdg.uis.unesco.org/2019/08/26/learningcounts-support-sdg-indicator-4-1-1-to-keep-learning-in-the-2030-agenda/ *

Montoya, S., & Tay-Lim, B. (2018). *Options in achieving global comparability for reporting on SDG 4* (CASGE Working Paper No. 3). Phoeniz, AZ: Arizona State University. *

Montoya, S., Beeharry, G. & Woolf, E. (2019, January 22). *A partnership for a global public good: data to improve learning* (Blog post – UIS). Retrieved from: https://sdg.uis.unesco.org/2019/01/22/a-partnership-for-a-global-public-good-data-to-improve-learning/ *

Moriarty, K. (2019). *Developing a transformative vision of global education? Unpacking education quality and learning in the policy formulation and content of Sustainable Development Goal 4* (Doctoral dissertation, University of Sussex). Retrieved from: http://sro.sussex.ac.uk/id/eprint/85462/

Morris, Z. S. (2009). The truth about interviewing elites. *Politics*, *29*(3), 209–217.

Motivans, A. (2015, June 5). *Building consensus towards an indicator framework to monitor education beyond 2015* (Blog post – NORRAG). Retrieved from: https://www.norrag.org/building-consensus-towards-an-indicator-framework-to-monitor-education-beyond-2015/ *

Mullis, I. V. S., & Martin, M. O. (2015). *Introducing IEA's LaNA for Developing Countries. P*resentation delivered in the context of the 56th IEA General Assembly. Retrieved from:

https://www.iea.nl/publications/presentations/ga56introducing-ieas-lana-developing-countries

Müller, H., Fey, M., & Rauch, C. (2013). Winds of change: exogenous events and trends as norm triggers (or norm killers). In H. Müller, & C. Wunderlich (Eds.), *Norm Dynamics in Multilateral Arms Control. Interests, Conflicts and Justice* (pp. 141–160). Athens, GA: The University of Georgia Press.

Mundy, K. (1998). Educational multilateralism and world (dis)order. Comparative Education Eeview, 42(4), 448–478.

Mundy, K. (1999). Educational multilateralism in a changing world order: UNESCO and the limits of the possible. *International Journal of Educational Development, 19*(1), 27–52.

Mundy, K. (2006). Education for all and the new development compact. *Review of Education, 53*, 23–48.

Mundy, K. (2010). Education for All and the global governors. In D. D'Avant, M. Finnemore, & S. K. Sell (Eds.), *Who Governs the Globe?* (pp. 333–355). Cambridge: Cambridge University Press.

Mundy, K. (2016). "Leaning in" on Education for All. CIES Presidential Address. *Comparative Education Review, 60*(1), 1–26.

Mundy, K., & Manion, C. (2015). The Education for All Initiative: History and Prospects Post-2015. In T. McCowan, & E. Unterhalter (Eds.), *Education and International Development. An Introduction* (pp. 49–68). London/New York: Bloomsbury.

Mundy, K., & Murphy, L. (2001). Transnational advocacy, global civil society? Emerging evidence from the field of education. *Comparative Education Review, 45*(1), 85–126.

Murphy, H., & Kellow, A. (2013). Forum shopping in global governance: understanding states, business and NGOs in multiple arenas. *Global Policy, 4*(2), 139–149.

Nader, L. (1972). Up the anthropologist – Perspectives gained from studying up. In D. Hymes (Ed.), *Reinventing Anthropology* (pp. 284–311). New York: Pantheon Books.

Nelson, S. C. (2014). Playing favorites: how shared beliefs shape the IMF's lending decisions. *International Organization 68*(2), 297–328.

Nelson, S. C., & Weaver, C. (2016). Organizational culture. In J. K. Cogan, I. Hurd, & J. Johnstone (Eds*.), The Oxford Handbook of International Organizations* (pp. 920–939). Oxford: Oxford University Press.

Neumann, I. B. (2002). Returning practice to the linguistic turn: the case of diplomacy. *Millennium: Journal of International Studies, 31*(3), 627–651.

Nicolai, S., Wild, L., Wales, J., Hine, S., & Engel, J. (2014). *Unbalanced Progress. What Political Dynamics Mean for Education Access and Quality* (ODI Working Paper No. 5). London: Overseas Development Institute. *

O'Connor, H., & Madge, C. (2017). Online interviewing. In N. F. Fielding, R. M. Lee, & G. Blank (Eds.), *The SAGE Handbook of Online Research Methods* (pp. 416–434). London: SAGE.

OECD. (1996). *Shaping the 21st Century: The Contribution of Development Co-operation.* Paris: OECD.

OWG. (2013a). *Co-Chairs' Summary bullet points from OWG-4.* Retrieved from: https://sustainabledevelopment.un.org/content/documents/1872mgsummary.pdf *

OWG. (2013b). *Concluding remarks of Co-Chairs, OWG4 19 June 2013*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/3693cochairsconcluding.pdf *

OWG. (2013c). *Co-Chair's summary on meetings with the Major Groups and other Stakeholders*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/1872mgsummary.pdf *

OWG. (2014a). *Focus Area Document*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/3276focusareas.pdf *

OWG. (2014b). *Compendium of existing goals and targets under the 19 Focus Areas being considered by the Open Working Group*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/3507Existing%20targets_1_April_version.xlsx *

OWG. (2014c). *Encyclopedia Groupinica: A compilation of goals and targets. Suggestions from OWG-10.* Retrieved from: https://sustainabledevelopment.un.org/content/documents/3698EncyclopediaGroupinica.pdf *

OWG. (2014d). *Working Document for 5-9 May Session of Open Working Group – 11th session of the Open Working Group (May 2014)*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/3686WorkingDoc_0205_additionalsupporters.pdf *

OWG. (2014e). *Introduction and proposed goals and targets on sustainable development for the post-2105 development agenda – 12th session of the Open Working Group (June 2014)*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/4528zerodraft12OWG.pdf *

OWG. (2014f). *Introduction to the Proposal of the Open Working Group for Sustainable Development Goals – 13th session of the Open Working Group (July 2014)*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/4523zerodraft.pdf *

Ozga, J. (2009). Governing education through data in England: from regulation to self-evaluation. *Journal of Education Policy, 24*(2), 149–162.

Packer, S. (2007). *International EFA Architecture Lessons and Prospects: A Preliminary Assessment* (Background paper prepared for the Education for All Global Monitoring Report 2008). Paris: EFA-GMR/UNESCO.

PARIS21. (2015). *A Road Map for a Country-led Data Revolution*. Paris: OECD.

Park, S. (2005). Norm diffusion within international organizations: a case study of the World Bank. *Journal of International Relations and Development, 8*(2), 111–141.

Park, S. (2006). Theorizing norm diffusion within international organizations. *International Politics, 43*(3), 342–361.

Park, S., & Vetterlein, A. (2010). Do policy norms reconstitute global development? In S. Park & A. Vetterlein (Eds.), *Owning Development. Creating Policy Norms in the IMF and the World Bank* (pp. 225–247). Cambridge/New York: Cambridge University Press.

Patton, M. Q. (2015). *Qualitative Research and Evaluation Methods: Integrating Theory and Practice* (4th edition). Thousand Oaks, CA: SAGE.

Payne, R. (2001). Persuasion, frames and norm construction. *European Journal of International Relations 7*(1), 37–61.

Porter, T. M. (1995). *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life.* Princeton, NJ/West Sussex, UK: Princeton University Press.

Pouliot, V., & Mérand, F. (2013). Bourdieu's concepts. In R. Adler-Nissen (Ed.), *Bourdieu in International Relations. Rethinking Key Concepts in IR* (pp. 24–44). Abingdon, UK/New York: Routledge.

Power, M. (1999). *The Audit Society: Rituals for Verification*. Oxford: Oxford University Press.

Puryear, J. M. (1995). International education statistics and research: status and problems. *International Journal of Educational Development, 15*(1), 79–91.

Rao, V., & Woolcock, M. (2007). The disciplinary monopoly in development research at the World Bank. *Global Governance, 13*(4), 479–84.

Reckwitz, A. (2002). Toward a theory of social practices: a development in culturalist theorizing. *European Journal of Social Theory, 5*(2), 243–263.

Reddy, S. G., & Kvangraven, I. H. (2015). *Global development goals: if at all, why, when and how?* (Working document). Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2666321

Risse, T. (2004). Global governance and communicative action. *Government and Opposition, 39*(2), 288–313.

Risse, T., Ropp. S., & Sikkink, K. (Eds.) (1999). *The Power of Human Rights: International Norms and Domestic Change.* New York: Cambridge University Press.

Ritchie, J., Lewis, J., & Elam, G. (2003). Designing and selecting samples. In J. Ritchie & J. Lewis (Eds.), *Qualitative Research Practice. A Guide for Social Science Students and Researchers* (pp. 77–108). London: SAGE.

Roberts, J. (2005). Millennium Development Goals: Are international targets now more credible? *Journal of International Development, 17*(1), 113–129.

Robinson, C. (2014, September 4). *A personal perspective on the follow up after Dakar* (Blog post – World Education Blog). Retrieved from: https://gemreportunesco.wordpress.com/2014/09/04/a-personal-perspective-on-the-follow-up-after-dakar/

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

Rose, N. (1999). *Powers of Freedom: Reframing Political Thought.* Cambridge: Cambridge University Press.

Rose, P. (2005). Is there a "fast-track" to achieving education for all? *International Journal of Educational Development, 25*(4), 381–394.

Ryle, G. (1949) *The Concept of Mind.* Hutchinson: London.

Sachs-Israel, M. (2017). The SDG4-Education 2030 Agenda and its Framework For Action – The process of its development and first steps in taking it forward. *Bildung und Erziehung, 69*(3), 269–290.

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

Salmons, J. (2015). *Qualitative Online Interviews* (2nd edition). Thousand Oaks, CA: SAGE.

Sandefur, J., & Glassman, A. (2014). *The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics* (CGD Working Paper No. 373). Washington, D.C.: Center for Global Development.

Sandholtz, W. (2008) Dynamics of international norm change: rules against wartime plunder. *European Journal of International Relations 14*(1), 101–131.

Save the Children. (2013). *Ending the Hidden Exclusion. Learning and Equity in Education post-2015.* London: Save the Children. *

Save the Children. (2015). *Results for Children. 2015 Annual Review*. Fairfield, CT/Washington, D.C.: Save the Children. *

Sayed, Y., & Ahmed, R. (2015). Education quality, and teaching and learning in the post-2015 education agenda. *International Journal of Educational Development, 40*, 330–338.

Sayed, Y., & Moriarty, K. (2020). SDG4 and the "Education quality turn": Prospects, possibilities and problems. In A. Wulff (Ed.), *Tensions, Threats and Opportunities in the Sustainable Development Goal on Quality Education* (pp. 194–213). Leiden, The Netherlands/Boston, MA: Brill | Sense.

Sayed, Y., Ahmed, R., & Mogliacci, R. (2018). The 2030 Global Education Agenda and the SDGs: process, policy and prospects. In A. Verger, M. Novelli, & H. K. Altinyelken (Eds.), *Global Education Policy and International Development* (2nd edition) (pp. 185–208). London/New York: Bloomsbury.

Schatz, E. (2009). Ethnographic immersion and the study of politics. In E. Schatz (Ed.), *Political Ethnography. What Immersion Contributes to the Study* (pp. 1–22). Chicago, IL/London: The Univeristy of Chicago Press.

Schatzki, T., K. K. Cetina, & von Savigny, E. (Eds.). (2001). *The Practice Turn in Contemporary Theory*. London: Routledge.

Schroeder, M. B. (2014). Executive leadership in the study of international organization: a framework for analysis. *International Studies Review, 16*(3), 339–361.

Schweisfurth, M. (2015). Learner-centred pedagogy: towards a post-2015 agenda for teaching and learning. *International Journal of Educational Development, 40*, 259–266.

SDSN. (2013). *An Action Agenda for Sustainable Development. Report for the Un Secretary-General*. New York: Sustainable Development Solutions Network. *

SDSN. (2014). *The Future of Our Children: Lifelong, Multi-Generational Learning for Sustainable Development. Report prepared by Thematic Group 4 (Early Childhood Development, Education and Transition to Work).* New York: https://resources.unsdsn.org/the-future-of-our-children-lifelong-multi-generational-learning-for-sustainable-development *

SDSN. (2015). *Indicators and Monitoring Framework for Sustainable Development Goals. Launching a Data Revolution for the SDG*. New York: SDSN. *

Seabrooke, L., & Henriksen, L. F. (2017). Issue control in transnational professional and organizational networks. In L. Seabrooke, & L. F. Henriksen (Eds.), *Professional Networks in Transnational Governance* (pp. 3–24). Cambridge: Cambidge University Press.

Seabrooke, L., & Sending, O. J. (2015). *Open systems of international organization* (GR: EEN Working Paper No. 51). Coventry, UK: Centre for the Study of Globalisation and Regionalisation, Department of Politics and International Studies, University of Warwick.

Seabrooke, L., & Tsingou, E. (2009). *Revolving doors and linked ecologies in the world economy: policy locations and the practice of international financial reform* (CSGR Working Paper No. 260/09 ). Coventry, UK: Centre for the Study of Globalisation and Regionalisation, Department of Politics and International Studies, University of Warwick.

Sedgwick, M., & Spiers, J. (2009). The use of videoconferencing as a medium for the qualitative interview. *International Journal of Qualitative Methods, 8*(1), 1–11

Sen, A. (1985) *Commodities and Capabilities*. Amsterdam: North-Holland.

Sending, O. J. (2002). Constitution, choice and change: problems with the "logic of appropriateness" and its use in constructivist theory. *European Journal of International Relations, 8*(4), 443–70.

Sending, O. J. (2015). *The Politics of Expertise: Competing for Authority in Global Governance*. Ann Arbor, MI: University of Michigan Press.

Sending, O. J., & Neumann, I. B. (2011). Banking on power: how some practices in an international organization anchor others. *Cambidge Studies in International Relations, 119*(1), 231–254.

Sending, O. J., Pouliot, V., & Neumann, I. B. (2015). Introduction. In O. J. Sending, V. Pouliot, & I. B. Neumann, (Eds.), *Diplomacy and the Making of World Politics* (pp. 1–28). Cambridge: Cambridge University Press.

Sénit, C. A. (2020). Leaving no one behind? The influence of civil society participation on the Sustainable Development Goals. *Environment and Planning C: Politics and Space*, *38*(4), 693–712.

Sénit, C. A., Biermann, F., & Kalfagianni, A. (2017). The representativeness of global deliberation: a critical assessment of civil society consultations for sustainable development. *Global Policy*, *8*(1), 62–72.

Sexsmith, K., & McMichael, P. (2015). Formulating the SDGs: reproducing or Reimagining State-Centered Development? *Globalizations, 12*(4), 581–596.

Shore, C., & Wright, S. (2015). Governing by numbers: audit culture, rankings and the new world order. *Social Anthropology, 23*(1), 22–28.

Siddiqui, F. R., Friedman, E. A., & Nader, A. (2014). Representativeness of the UN post-2015 national consultations. *The Lancet Global Health*, *2*(11), e631–e632.

Smith, K. E. (2006). Problematising power relations in "elite" interviews. *Geoforum, 37*(4), 643–553.

Spicer, N., Agyepong, I., Ottersen, T., Jahn, A., & Ooms, G. (2020). "It's far too complicated": why fragmentation persists in global health. *Globalization and Health, 16.* Online publication. doi: 10.1186/s12992-020-00592-1

Spijkers, O., & Honniball, A. (2015). Developing global public participation (2): shaping the Sustainable Development Goals. *International Community Law Review*, *17*(3), 251–296.

Stanton, E. A. (2007). *The Human Development Index: A History* (PERI Working Paper Series No. 127). Amherst, MA: Political Economy Research Institute (PERI) – University of Massachusetts-Amherst.

Swartz, D. L. (2013). Metraprinciples for sociological research in a Bourdieusian perpective. In P. S. Gorski (Ed.), *Bourdieu and Historical Analysis* (pp. 19–35). Durham/London: Duke University Press.

Takala, T. & Piattoeva, N. (2017, May 23). *Russia and education assistance: searching for its role as a returning donor* (Blog post – NORRAG). Retrieved from: https://www.norrag.org/russia-education-assistance-searching-role-returning-donor/

Tallberg, J., Lundgren, M., Sommerer, T., & Squatrito, T. (2018. February). *Norm adoption by International Organizations, 1980-2015*. Paper presented at the Annual Conference on The Political Economy of International Organizations, Madison, WI. Retrieved from: https://www.peio.me/wp-content/uploads/2018/02/TallbergLundgrenSommererSquatrito-PEIO-paper-43.pdf

Tansey, O. (2007). Process tracing and elite interviewing: a case for non-probability sampling. *PS: Political Science and Politics*, *40*(4), 765–772.

Taylor, M. (2016). *The Political Economy of Statistical Capacity* (Discussion Paper No. IDB-DP-471). Washington, D.C.: Inter-American Development Bank.

Thérien, J. P., & Pouliot, V. (2020). Global governance as patchwork: the making of the Sustainable Development Goals. *Review of International Political Economy*, *27*(3), 612–636.

Thomas, G., Meyer, J. W., Ramirez, F., & Boli, J. (1987). *Institutional Structure*. Newbury Park, CA: SAGE.

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.

Tikly, L. (2017). The future of Education for All as a global regime of educational governance. *Comparative Education Review, 61*(1), 22–57.

Tobin, M., Lietz, P., Nugroho, D., Vivekananan, R., & Nyamkhuu, T. (2015). Using Large-Scale Assessments of Students' Learning to Inform Education Policy: Insights from the Asia-Pacific Region. Melbourne/Bangkok: ACER/UNESCO Bangkok Office.

Torres, R. M. (2000). *One Decade of Education for All: The Challenge Ahead*. Paris: IIEP-UNESCO.

Torres, R. M. (2001). What happened at the World Education Forum? *Adult Education and Development, 56.* Online publication. Retrieved from: https://www.dvv-international.de/en/adult-education-and-development/editions/aed-562001/literacy-and-basic-education/what-happened-at-the-world-education-forum/#:~:text=The%20purpose%20of%20the%20Forum,order%20to%20continue%20the%20task.

Tota, P. M. (2014). Filling the gaps: the role and impact of international non-governmental organisations in "Education for All". *Globalisation, Societies and Education, 12*(1), 92–109.

Treviño, E., & Órdenes, M. (2017). *Exploring Commonalities and Differences in Regional and International Assessments* (UIS Information Paper No. 48). Montreal, Canada: UIS.

UIS. (2008). *Report on the activities of the Institute in 2008.* Montreal, Canada: UIS. *

UIS. (2009). *Report on the activities of the Institute in 2009*. Montreal, Canada: UIS. *

UIS. (2011). *Report on the activities of the Institute in 2011*. Montreal, Canada: UIS. *

UIS. (2014). *Learning Outcomes Advisory Board. Terms of Reference*. Montreal, Canada: UIS. *

UIS. (2015). *Catalogue of Learning Assessments. A public resource to learning assessments around the world.* Montreal, Canada: UIS. Retrieved from: http://uis.unesco.org/sites/default/files/documents/catalogue-of-learning-assessments-brochure-education-2015-en.pdf *

UIS. (2016a). *Proposal to develop global learning metrics. How to measure SDG 4.1.* Montreal, Canada: UIS. Retrieved from: http://uis.unesco.org/sites/default/files/documents/proposal-to-develop-global-learning-metrics.pdf *

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

UIS. (2016c). *Global Alliance to Monitor Learning. Concept note*. Montreal, Canada: UIS. Retrieved from: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/10/gaml-concept_paper-2017-en2_0.pdf *

UIS. (2016d). *Global Alliance to Monitor Learning. List of participants*. Montreal, Canada: UIS. Retrieved from: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/08/Confirmed-participants_GAML2-meeting_Website_20161101.pdf *

UIS. (2016e). *Consensus and chartering the way forward.* Presentation delivered in the context of the GAML 2nd Meeting. Retrieved from: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/08/Session-15_Consensus-and-chartering-way-forward.pdf *

UIS. (2016f). *Implementation strategy for measuring SDG4 learning outcomes targets.* Montreal, Canada: UIS. Retrieved from: http://uis.unesco.org/sites/default/files/documents/implementation-strategy-for-measuring-sdg4-learning-outcomes-targets-2016-en.pdf *

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

UIS. (2017b). *Counting the number of Children Not Learning. Methodology for a global composite indicator for education* (UIS Information Paper No. 47). Montreal, Canada: UIS. *

UIS. (2017c). *The Global Alliance to Monitor Learning (GAML): Governance and organization*. Montreal, Canada: UIS. Retrieved from: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/10/gaml-governance-organization-2017-en.pdf *

UIS. (2017d). *Cross-national assessments technical meeting (September 2017, Hamburg): Summary.* Montreal, Canada: UIS. Retrieved from: http://uis.unesco.org/sites/default/files/documents/gaml-cross-national-assessments-technical-meeting-summary-hmaburg-2017-en.pdf *

UIS. (2018a). *Global Alliance to Monitor Learning (GAML): 2018 Progress report*. Montreal, Canada. UIS. Retrieved from: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/10/GAML-5-Report.pdf *

UIS. (2018b). *SDG 4 Reporting: linking to the UIS Reporting Scale through social moderation.* Montreal, Canada: UIS. Retrieved from: http://uis.unesco.org/sites/default/files/documents/gaml4-sdg4-reporting-linking-uis-reporting-scale-social-moderation.pdf *

UIS. (2018c). *Costs and Benefits of Different Approaches to Measuring the Learning Proficiency of Students (SDG Indicator 4.1.1)* (UIS Information paper No. 3 - UIS/2018/ED/IP/53). Montreal, Canada: UIS. *

UIS. (2018d). *FAQ for the application to upgrade SDG indicator 4.1.1(a).* Montreal, Canada: UIS. Retrieved from: https://unstats.un.org/sdgs/files/meetings/iaeg-sdgs-meeting-08/4.5a%20UNESCO-UIS%204.1.1a%20FAQs.pdf *

UIS. (2019a). *SDG 4 Data Digest How to Produce and Use the Global and Thematic Education Indicators.* Montreal, Canada: UIS. *

UIS. (2019b). *Global Alliance to Monitor Learning: Update on progress.* Presentation delivered by the UIS Director in the context of the 6th GAML meeting. Retrieved from: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2019/05/GAML6-Session1-Progress.pdf *

UIS. (2020). SDG 4.1.1 *Measurement tools – Applications in the wake of COVID-19.* Montreal, Canada: UIS. Retrieved from: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2020/06/Presentation-1-GAML-Webinar_GPF-and-Policy-Linking_UIS.pdf *

UIS/ACER. (2014). *Learning Metrics Partnership.* Montreal, Canada/Melbourne, Australia: UIS, Australian Aid, DFAT and ACER. Retrieved from http://uis.unesco.org/sites/default/files/documents/learning- metrics-partnership-a-capacity-support-and-policy-strengthening-initiative-to- develop-and-use-common-learning-metrics-mathematics-reading-2014-en_1.pdf *

UIS/CUE. (2013a). *Toward Universal Learning. What Every Child Should Learn*. Montreal, Canada/Washington, D.C.: UIS/Center for Universal Education at Brookings. *

UIS/CUE. (2013b). *Toward Universal Learning. Recommendations from the Learning Metrics Task Force.* Montreal, Canada/Washington, D.C.: UIS/Center for Universal Education at Brookings. *

UN News. (2014, August 24). *Data revolution advisory group named by UN Secretary-General.* Retrieved from: https://news.un.org/en/story/2014/08/476242-data-revolution-advisory-group-named-un-secretary-general

UN Secretary-General. (2010). *Keeping the promise. A forward-looking review to promote an agreed action agenda to achieve the Millennium Development Goals by 2015* (A/64/665). New York: United Nations General Assembly. *

UN Secretary-General. (2013). *A Life of Dignity for All: Accelerating Progress Towards the Millennium Development Goals and Advancing the United Nations*

*Development Agenda Beyond 2015 – Report of the Secretary-General* (A/68/202). New York: United Nations General Assembly. *

UN. (2012). *Terms of Reference for the High-level Panel of Eminent Persons on the Post-2015 Development Agenda. Retrieved* from: https://www.un.org/sg/sites/www.un.org.sg/files/documents/management/ToRpost2015.pdf *

UN. (2013). *A New Global Partnership: Eradicate Poverty and Transform Economies through Sustainable Development. The Report of the High-Level Panel of Eminent Persons on the Post-2015 Development Agenda*. New York: United Nations. *

UN. (2014). *We the Peoples. Celebrating 7 Million Voices*. New York: United Nations. *

UN. (2015a). *Summary of Stock taking meeting of Intergovernmental Negotiations on the Post-2015 Development Agenda. 21 January 2015 - Remarks delivered by Ambassador David Donoghue, Permanent Representative of Ireland to the UN*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/5913Summary%20of%20IGN%20Stock%20Taking%20Mtg%2019_21%20Jan%202015.pdf *

UN. (2015b). *Addis Ababa Action Agenda of the Third International Conference on Financing for Development (Addis Ababa Action Agenda).* New York: United Nations. *

UN. (2020). *Participate with the Open Working Group on SDGs.* Retrieved from: https://sustainabledevelopment.un.org/majorgroups/post2015/thematicclusters/participate

UN/DESA. (2020). *Introduction to ECOSOC Consultative Status*. Retrieved from: https://csonet.org/?menu=30

UNDG. (2013a). *Summary note on Post-2015 Consultations prepared by UNDP for the OWG.* Retrieved from: https://sustainabledevelopment.un.org/content/documents/1727undp.pdf *

UNDG. (2013b). *A Million Voices. The World We Want. A Sustainable Future with Dignity for All. N*ew York: UN Development Group. *

UNDG. (2013c). *The Global Conversation Begins. Emerging Views for a New Development Agenda. N*ew York: UN Development Group. *

UNECE. (2015). *Update on the ongoing process to develop an indicator framework for the Post-2015 Development Agenda. Note by the United Nations Statistical Division* (ECE/CES/2015/36). Geneva:  UNECE. *

UNESCO Internal Oversight Service. (2016). *Evaluation of the Education for All (EFA) Global and Regional Coordination Mechanisms* (IOS/EVS/PI/150). Paris: UNESCO.

UNESCO Internal Oversight Service. (2019). *Evaluation of the Future of UNESCO's Education Sector: The Normative vs Operational Role in the Context of Agenda 2030* (IOS/EVS/PI 184). Paris: UNESCO.

UNESCO News. (2015, July 8). *Director-General welcomes new Commission on Financing of Global Education.* Retrieved from: https://en.unesco.org/themes/110855/news?order=created&sort=asc&page=48

UNESCO. (2011). *Decisions adopted by the Executive Board at its 186ᵗʰ Session* (186 EX/Decision). Paris: UNESCO. *

UNESCO. (2012). *Education for All (EFA) Part 1. Strategic Vision and Plan for UNESCO's Advocacy Efforts for EFA* (190 EX/7). Paris: UNESCO. *

UNESCO. (2013a). *Records of the General Conference, 37ᵗʰ session* (Volume 1). Paris: UNESCO. *

UNESCO. (2013b). *Report by the director-general on the execution of the programme adopted by the general conference.* Paris: UNESCO. *

UNESCO. (2014a). *Position Paper on Education Post-2015* (ED-14/EFA/POST-2015/1). Paris: UNESCO. *

UNESCO. (2014b). *Education Beyond 2015* (195 EX/6). Paris: UNESCO. *

UNESCO. (2014c). *Joint Proposal of the EFA Steering Committee on Education Post-2015* (ED-14/EFA/POST-2015/2 Rev). Paris. UNESCO. *

UNESCO. (2014d). *2014-2021 Medium-Term Strategy* (37 C/4). Paris: UNESCO.

UNESCO. (2014e). *Global Education for All Meeting. 2014 GEM Final Statement – The Muscat Agreement* (ED-14/EFA/ME/3). Paris: UNESCO. *

UNESCO. (2014f). *SISTER Online Report: 37 C/5 - Monitoring of Programme Implementation for Regular Programme, Extrabudgetary and Category 1 Institutes and Centres Resources as at 30/06/2014.* Paris: UNESCO. *

UNESCO. (2015a). *Education for All Steering Committee. List of members.* Retrieved from: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/ED_new/pdf/EFA-SC_Members-2015.pdf

UNESCO. (2015b). *Outcomes of the World Education Forum 2015* (197 EX/6). Paris: UNESCO. *

UNESCO. (2015c). *UNESCO's participation in the preparations for a Post-2015 Development Agenda* (197 EX/7). Paris: UNESCO. *

UNESCO. (2016a). *First Meeting of the SDG-Education 2030 Steering Committee, 25-26 May 2016* (ED-2016/ED2030/ME/1). Paris: UNESCO. *

UNESCO. (2016b). *Meeting of the Collective Consultation of NGO on Education for All Coordination Group 23-24 May 2016 – Report.* Retrieved from: http://www.arabcampaignforeducation.org/functions.php?action=files&table=files&ID=328 *

UNESCO. (2018). *Global Education Meeting 2018. Synthesis Report* (ED-2018/ME-GEM). Paris: UNESCO. *

UNESCO. (2019). *The Promise of Large-Scale Learning Assessments. Acknowledging Limits to Unlock Opportunities*. Paris: UNESCO.

UNESCO/UIS. (2015). *Thematic indicators to monitor the Education 2030 Agenda. Technical Advisory Group proposal – October 2015*. Retrieved from: https://unesdoc.unesco.org/ark:/48223/pf0000235172?posInSet=1&queryId=fdff71e8-5a95-4668-b048-c56860d07b0b *

UNGA. (2001). *Road map towards the implementation of the United Nations Millennium Declaration. Report of the Secretary-General* (A756/326). New York: UN General Assembly. *

UNGA. (2010). *Resolution adopted by the General Assembly on 22 September 2010* (A/RES/65/1). New York: UN General Assembly. *

UNGA. (2013a). *General Assembly Open Working Group on Sustainable Development Goals. Methods of work.* Retrieved from: https://sustainabledevelopment.un.org/content/documents/1692OWG_methods_work_adopted_1403.pdf *

UNGA. (2013b). *Summary of the First Meeting of the Open Working Group of the General Assembly on Sustainable Development Goals. 14-15 March 2013*. Retrieved from: https://sustainabledevelopment.un.org/content/documents/1700summaryowg.pdf *

UNGA. (2013c). *Resolution adopted by the General Assembly on 9 October 2013. 68/6. Outcome document of the special event to follow up efforts made towards achieving the Millennium Development Goals* (A/RES/68/6). New York: United Nations General Assembly. *

UNGA. (2014a). *A Compendium of Issues Briefs Prepared by the United Nations Inter-Agency Technical Support Team for the United Nations General Assembly Open Working Group on Sustainable Development Goals.* New York: UN/DESA. *

UNGA. (2014b). *The road to dignity by 2030: Ending Poverty, Transforming All Lives and Protecting the Planet. Synthesis Report of the Secretary-General on the Post-2015 Sustainable Development Agenda* (A/69/700). New York: UN General Assembly. *

UNGA. (2015). *Resolution adopted by the General Assembly on 25 September 2015. 70/1. Transforming our world: the 2030 Agenda for Sustainable Development* (A/RES/70/1). New York: United Nations General Assembly. *

UNGA. (2017). *Resolution adopted by the General Assembly on 6 July 2017 - 71/313. Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development* (A/RES/71/313). New York:  United Nations General Assembly. *

UNICEF. (2013). *UNICEF Key Messages on the Post-2015 Development Agenda*. New York: UNICEF. *

UNICEF. (2019). *Every Child Learns. UNICEF Education Strategy 2019-2030.* New York: UNICEF. *

UNICEF/UNESCO. (2013). *Envisioning Education in the Post-2015 Development Agenda. Global Thematic Consultation in the Post-2015 Development Agenda.* New York/Paris: UNICEF/UNESCO. *

Universalia Management Group. (2010). *Evaluation of the Global Monitoring Report – Final Report (Volume 1).* Montreal, Canada: Universalia. Retrieved from: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/GMR/pdf/diverse/gmr-evaluation-2009-vol1.pdf

UNSC. (2020a). *Tier classification for global SDG indicators*. Retrieved from: https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/

UNSC. (2020b). *IAEG-SDGs 2020 Comprehensive Review Process*. Retrieved from: https://unstats.un.org/sdgs/iaeg-sdgs/2020-comp-rev/

UNSC. (2021). *United Nations Statistical Commission. Overview*. Retrieved from: https://unstats.un.org/unsd/statcom. *

Unterhalter, E. (2007). *Gender, Schooling and Global Social Justice.* London/New York: Routledge.

Unterhalter, E. (2014). Walking backwards into the future: a comparative perspective on education and a post-2015 framework. *Compare: A Journal of Comparative and International Education, 44*(6), 852–873.

Unterhalter, E. (2019). The many meanings of quality education: politics of targets and indicators in SDG4. *Global Policy, 10*(S1), 39–51.

Unterhalter, E., Poole, N., & Winters, N. (2015). Education, information, and knowledge. In J. Waage, & C. Yap (Eds.), *Thinking Beyond Sectors for Sustainable Development* (pp. 51–61). London: Ubiquity Press.

UNTT. (2012a). *Review of the contributions of the MDG Agenda to foster development: Lessons for the post-2015 UN development agenda. Discussion Note*. New York: UN/DESA. *

UNTT. (2012b). *Realizing the Future We Want for All. Report to the Secretary-General*. New York: UN/DESA. *

UNTT. (2012c). *UN System Task Team to support the preparation of the Post-2015 UN Development Agenda. Draft Concept Note, 6th January 2012*. Retrieved from: https://www.un.org/en/development/desa/policy/untaskteam_undf/concept_note.pdf *

UNTT. (2013). *A Renewed Global Partnership for Development*. New York: UN/DESA. *

USAID. (2011). *Education. Opportunity through Learning. USAID Education Strategy.* Washington, D.C.: USAID. *

Vandemoortele, J. (2008). Making sense of the MDGs. *Development, 51*(2), 220–227.

Vandemoortele, J. (2011). The MDG story: intention denied. *Development and Change, 42*(1), 1–21.

Varghese, N. V. (Ed.). (2012). *From Schooling to Learning. A Report from the IWGE*. Paris: UNESCO International Institute for Educational Planning. *

Vauchez, A. (2011). Interstitial power in fields of limited statehood: introducing a "weak field" approach to the study of transnational settings. *International Political Sociology, 5*(3), 340345.

Verger, A., & Novelli, M. (Coords.). (2012). *Campaigning for "Education for All" Histories, Strategies and Outcomes of Transnational Advocacy Coalitions in Education.* Rotterdam, The Netherlands: Sense Publishers.

Vetterlein, A., & Moschella, M. (2014). International organizations and organizational fields: explaining policy change in the IMF. *European Political Science Review, 6*(1), 143–165.

VVAA. (2015a). *Incheon Declaration. Education 2030: Towards inclusive and equitable quality education and lifelong learning for all* (ED/WEF2015/MD/3). Paris: UNESCO. *

VVAA. (2015b). *Education 2030. Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4* (ED-2016/WS/28). Paris: UNESCO. *

Waage, J., Banerji, R., Campbell, O., Chirwa, E., Collender, G., Dieltiens, V., Dorward, A., Godfrey-Faussett, P., Hanvoravongchai, P., Kingdon, G., Little, A., Mills, A., Mulholland, K., Mwinga, A., North, A., Patcharanarumol, C. P., Tangcharoensathien, V., & Unterhalter, E. (2010). The Millennium Development Goals: a cross-sectoral analysis and principles for goal setting after 2015: Lancet and London International Development Centre Commission. *The Lancet, 376*(9745), 991–1023.

Wagner, D. A. (Ed.). (2014). *Learning and Education in Developing Countries. Research and Policy for the Post-2015 UN Development Goals*. New York: Palgrave Macmillan. *

Wagner, D.A., Wolf, S., & Boruch, R.F. (Eds.) 2018. *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries*. Paris: UNESCO-IIEP.

Waldner, D. (2012). Process tracing and causal mechanisms. In H. Kincaid (Ed.), *The Oxford Handbook of Philosophy of the Social Science* (pp. 65–84). Oxford: Oxford University Press.

Ward, M. (2004). *Quantifying the World. UN Ideas and Statistics*. Bloomington, IN: Indiana University Press.

Weaver, C. (2008). *Hypocrisy Trap. The World Bank and the Poverty of Reform.* Princeton, NJ/Oxford: Princeton University Press.

Weaver, C. (2010). The strategic social construction of the World Bank's gender and development policy norm. In S. Park & A. Vetterlein (Eds.), *Owning Development. Creating Policy Norms in the IMF and the World Bank* (pp. 70-88). Cambridge/New York: Cambridge University Press.

Wendt, A. (1999). *Social Theory of International Politics*. New York: Cambridge University Press.

Wiener, A. (2003). Constructivism: the Limits of bridging gaps. *Journal of International Relations and Development, 6*(3), 252–275.

Wiener, A. (2007). Contested meanings of norms: a research framework. *Comparative European Politics, 5*(1), 1–17.

Wiener, A. (2009). Enacting meaning-in-use: qualitative research on norms and international relations. *Review of International Studies, 35*(1), 175–193.

Wiener, A. (2014). *A Theory of Contestation*. Heidelberg, Germany: Springer-Verlag Berlin Heidelberg.

Winthrop, R., & Anderson, K. (2013). Can international large-scale assessments inform a global learning goal? Insights from the learning metrics task force. *Research in Comparative and International Education, 8*(3), 279–293.

World Bank. (2011). *Learning for All: Investing in People's Knowledge and Skills to Promote Development. World Bank Education Sector Strategy 2020*. Washington, D.C.: The World Bank. *

World Bank. (2015). *Final READ Trust Fund Report 2008–2015* (Report No. 101527). Washington, D.C.: The World Bank. *

World Bank. (2017). *Learning to Realize Education's Promise.* Washington, D.C.: The World Bank. *

World Bank. (2019a). *Learning Poverty: What Will It Take?* Washington, D.C.: The World Bank. *

World Bank. (2019b). *Education Systems in Action: Transforming Data into Education Policy Implementation.* Washington, D.C.: The World Bank. *

World Bank. (2020). *The Human Capital Index 2020 Update. Human Capital in the Time of COVID-19.* Washington, D.C.: The World Bank. *

World Education Forum. (2000). *The Dakar Framework for Action. Education for All: Meeting our Collective Commitments.* Paris: UNESCO.

Wulff, A. (2020). The twists and turns in negotiating a global education goal: a civil society perspective. In A. Wulff (Ed.), *Tensions, Threats and Opportunities in the Sustainable Development Goal on Quality Education* (pp. 28–64). Leiden, The Netherlands/Boston, MA: Brill | Sense.

Wunderlich, C. (2013). Theoretical approaches to norm dynamics. In H. Müller, & C. Wunderlich (Eds.), *Norm Dynamics in Multilateral Arms Control. Interests, Conflicts and Justice* (pp. 20–47). Athens, GA: University of Georgia Press.

Yamada, S. (2016). Theorizing the paradigm shift in educational development. In S. Yamada (Ed.), *Post-Education-For All and Sustainable Development Paradigm: Structural Changes with Diversifying Actors and Norms* (pp. 1–34). Bingley, UK: Emerald.

Ydesen, C., & Grek, S. (2019). Securing organisational survival: a historical inquiry into the OECD's work in education during the 1960s. *Paedagogica Historica, 56*(3), 412–427.

Zuijderduijn, M., Egberts, F., & Krämer, E. (2016). *Outcome Evaluation of the UNDP Project "Building the Post-2015 Development Agenda: Open and Inclusive Consultations".* Ede, The Netherlands: MDF Training & Consultancy BV.

# Annexes

## Annex 1. Interview guides [Chapter 4]

### IG1. Interview guide relative to research objective 1

*Background data and individual engagement*

1. Please, could you tell me about your education and training, and your professional trajectory?

2. What is your specific position in the [*organization/institution/administration*] you represent? What are your job duties?

   >> Probe for: possible linkages with SDG4.

3. How have you been involved with the so-called post-2015 debate? What has been your role, contribution or input to the [*deliberation/ committee/ consultation/ working group mentioned by the interviewee*]? Why were you designated with this responsibility?

*Organizational engagement*

4. To what extent, and in what way, has your organization participated in the so-called post-2015 debate?

   >> If not spontaneously mentioned, probe for specific strands of the debate according to previous research on the individual profile.

   >> Only for organizations other than EFA conveners: how did you gain access to [*mentioned* fora *and strands of the debate*]?

   Probe for: invitation, formal representation in committees or consultation mechanisms.

5. When and how did you learn about the post-2015 process? What motivated the involvement of your organization? Why did you perceive this to be a relevant endeavor?

6. Did your organization/institution/the state you represent participate in the past edition of MDG/EFA goals?

   >> If so, how have the priorities/preferences of your organization changed?

*Organizational priorities*

7. What are the priorities or interests of your organization regarding the new education goals?

8. What determined this organizational position?

   >> Probe for: internal processes of consultation and debate, lobbying of specific states or constituencies, leadership and expertise of specific staffers, research pieces, etc.

9. Which strategies did you mobilize in order to advance your agenda?

10. Are your preferences reflected in the current formulation of the goals, targets and indicators?

11. To what extent would you say your organization has had a determining or influential role?

*Overview of the process*

12. In your opinion, which three events [*gathering, meeting, consultation*] have been the most relevant in the advancement of the debate? What moments would you say can be characterized as turning points?

>> For each event mentioned: What was decided or discussed? Why do you consider it to be a relevant moment? In what way did your organization participate? Which were the most determining organizations or individual actors in shaping the final outcome?

13. Can you identify any recurrent or punctual sources of disagreement, friction or tension?

>> If so, how has this been solved?

>> If so, would you say such disagreement predates the post-2015 debate? Would you say it outlived the process?

14. Would you say that the convergence between the EFA and the UN strands of the debate was a predictable outcome? What risks and opportunities would you say it entails?

15. In your view, what are the advantages and disadvantages of the architecture of the post-2015 debate?

>> Probe for: coexistence of multiple strands of debate, extensive use of consultation mechanisms, participation of CSO and NGO constituencies in UN negotiation chambers, division of labor between political and technical labor, etc.

>> Probe for: differences in relation to prior goal-setting exercises the respondent might be familiar with (MDGs, EFA).

*Partners and network*

16. With which organization/institution/represented states has your [*organization/ institution you represent*] been collaborating closely throughout the debate? In what ways have you been collaborating?

17. Which organizations or states would you say have played a particularly influential or prominent role throughout the debate? Why?

*Closing questions*

18. Is there anything else you think I might be interested in? Do you think I am missing any relevant debate or episode? Is there anything else you would like to discuss with me in relation to the issues we have been commenting on?

19. Could you kindly name a couple of individuals who, in your view, can provide relevant input to the questions we have been discussing?

**IG2. Interview guide relative to research objective 2**

*Background data and individual engagement*

1. Please, could you tell me about your education and training, and your professional trajectory?

2. What is your specific position in the [*organization/institution/administration*] you represent? What are your job duties?

3. How have you been involved in the efforts to define or produce the SDG4 indicators relative to learning? What has been your role, contribution or input in [*deliberation/ committee/ consultation/ working group mentioned by the interviewee*]? Why were you designated with this responsibility?

4. How did your organization get involved with the [*TCG, GAML, LMP or other initiatives oriented towards the production of comparable learning data*]?

   >> Probe for: invitation, personal acquaintances, engagement in previous initiatives, etc.

   >> What does this involvement entail? What kind of work do you perform, or what kind of input do you provide?

*A renewed emphasis on learning assessment*

5. In your judgment, which factors explain the growing emphasis on learning and assessment that characterizes the post-2015 or Education-2030 scenario? Can you think of particular events, organizations or publications that contributed to the increased attention given to learning and learning metrics?

6. In your view, what are the opportunities and limitations involved in the so-called quality turn - that is, the growing emphasis on learning outcomes, and the move away from an input-centered perspective?

7. Would you say that the new emphasis on learning metrics and the potential of large-scale assessments has impacted the agenda and organizational priorities of your organization?

8. How is your organization supporting the expansion, strengthening and use of learning data?

   >> Probe for: projects specific to different organizations (LT, Learning Poverty Indicator and HCI, NEQMAP, TALENT, Learning Portal, ANLAS, KI, etc.).

*The role of the UIS: agenda and network*

[Only for UIS staffers or informants with in-depth knowledge of the organization.]

9. How would you describe the evolution of the UIS role within the EFA/SDG4 architecture - given the fact that it has not usually enjoyed such a prominent role?

10. In your view, what are the organizational challenges faced by the Institute as a result of its new role and mandate?

11. Taking your time at UIS as a reference, in what ways would you say that the new emphasis on learning metrics impacted the Institute's agenda and organizational priorities?

    >> Probe for: specific programs and projects, as well as new divisions, teams or posts, created in order to support the use or production of large-scale assessment data.

12. Who were the UIS' closest collaborators (including international organizations and bilateral or multilateral agencies, UNESCO partners, etc.) in assessment-related projects or efforts? How or why were these partners selected?

    >> How has this network evolved during your years at UIS?

    >> If they have been longstanding partners of the Institute, in what ways would you say that the nature of the relationship differs from previous initiatives?

       Probe for: division of labor, nature of the expertise brought by each partner, etc.

13. In what ways would you say that the Institute's priorities or preferences in relation to large-scale assessments differ from the view put forward by other divisions (or institutes) within UNESCO?

*The work of GAML*

[Only for individuals directly engaged with the initiative.]

14. In your view, what are the main challenges involved in the development of a global framework for the assessment of learning process?

    >> How are they being addressed? Are you satisfied with this course of action?

15. Are you aware of any areas of contention that may have emerged during this process? Do you know how are they being resolved?

16. The monitoring of SDG4 learning targets requires the development, mapping and harmonization of a wide range of assessment modalities (including national assessments, cross-national assessments, citizen-led assessments, etc.). To your knowledge, is there any assessment modality that is being given priority in the efforts to develop global learning metrics?

17. Which assessment modalities do you think are more likely to be adopted by low- and middle-income countries in the near future?

    >> What role do you think might have been played by the technical work of the [*UIS/GPE/IIEP/World Bank – selected on the basis of the interviewee profile*] in this regard?

    >> In your view, is there any assessment modality more likely to be supported by donors (including bilateral and multilateral agencies, foundations, etc.)?

*Closing questions*

18. Is there anything else you would like to discuss with me in relation to the issues we have been commenting on?

19. Are there any specific publications you would recommend to me to consult?

**Annex 2: Overview of proposed goals and targets** [Chapter 5]

**Table A.1. Proposed targets and goals – EFA workstream, 2014**

| Document | Overarching goal | Targets |
|---|---|---|
| UNESCO's Position Paper (February 2014)<br><br>UNESCO (2014a) | Ensure equitable quality education and lifelong learning for all by 2030 | 1. All children participate in and complete a full cycle of free, compulsory and continuous quality basic education of at least 10 years, including one year of pre-primary education, leading to relevant and measurable learning outcomes based on national standards.<br>2. Increase transition to and completion of quality upper secondary education by x %, with all graduates demonstrating relevant learning outcomes based on national standards.<br>3. Tertiary education systems are expanded to allow qualified learners to access and complete studies leading to a certificate, diploma or degree.<br>4. All youth and adults achieve literacy, numeracy and other basic skills at a proficiency level necessary to fully participate in a given society and for further learning.<br>5. Increase by x % the proportion of youth (15-24 years) with relevant and recognized knowledge and skills, including professional, technical and vocational to access decent work.<br>6. Increase by x % participation in continuing adult education and training programmes, with recognition and validation of non-formal and informal learning.<br>7. Close the teachers' gap by recruiting adequate numbers of teachers who are well- trained, meet national standards and can effectively deliver relevant content, with emphasis on gender balance.<br>8. All youth (15-24 years) people and adults have opportunities to acquire – supported by safe, gender-responsive, inclusive learning environments – relevant knowledge and skills to ensure their personal fulfilment, contribute to peace and an equitable and sustainable world.<br>9. All countries progress towards allocating 6% of their GNP to education and 20% of their government budget to education, prioritizing groups most in need.<br>10. All donors progress towards allocating at least 20% of their ODA or its equivalent to education, prioritizing countries and groups most in need. |

| Joint Proposal of the EFA Steering Committee on Education Post-2014 (April 2014)<br><br>(UNESCO, 2104d) | Ensure equitable and inclusive quality education and lifelong learning for all by 2030 | 1. By 2030, at least x% of girls and boys are ready for primary school through participation in quality early childhood care and education, including at least one year of free and compulsory pre-primary education, with particular attention to gender equality and the most marginalized.<br>2. By 2030, all girls and boys complete free and compulsory quality basic education of at least 9 years and achieve relevant learning outcomes, with particular attention to gender equality and the most marginalized.<br>3. By 2030, all youth and at least x% of adults reach a proficiency level in literacy and numeracy sufficient to fully participate in society, with particular attention to girls and women and the most marginalized.<br>4. By 2030, at least x% of youth and y% of adults have the knowledge and skills for decent work and life through technical and vocational, upper secondary and tertiary education and training, with particular attention to gender equality and the most marginalized.<br>5. By 2030, all learners acquire knowledge, skills, values and attitudes to establish sustainable and peaceful societies, including through global citizenship education and education for sustainable development.<br>Input targets<br>6. By 2030, all governments ensure that all learners are taught by qualified, professionally-trained, motivated and well-supported teachers.<br>7. By 2030, all countries allocate at least 4-6% of their Gross Domestic Product (GDP) or at least 15-20% of their public expenditure to education, prioritizing groups most in need; and strengthen financial cooperation for education, prioritizing countries most in need. |
|---|---|---|
| Muscat Agreement (May 2014)<br><br>(UNESCO, 2014e) | Ensure equitable and inclusive quality education and lifelong learning for all by 2030 | 1. By 2030, at least x% of girls and boys are ready for primary school through participation in quality early childhood care and education, including at least one year of free and compulsory pre-primary education, with particular attention to gender equality and the most marginalized.<br>2. By 2030, all girls and boys complete free and compulsory quality basic education of at least 9 years and achieve relevant learning outcomes, with particular attention to gender equality and the most marginalized.<br>3. By 2030, all youth and at least x% of adults reach a proficiency level in literacy and numeracy sufficient to fully participate in society, with particular attention to girls and women and the most marginalized.<br>4. By 2030, at least x% of youth and y% of adults have the knowledge and skills for decent work and life through technical and vocational, upper secondary and tertiary education and training, with particular attention to gender equality and the most marginalized.<br>5. By 2030, all learners acquire knowledge, skills, values and attitudes to establish sustainable and peaceful societies, including through global citizenship education and education for sustainable development. |

6. By 2030, all governments ensure that all learners are taught by qualified, professionally-trained, motivated and well-supported teachers.
7. By 2030, all countries allocate at least 4-6% of their Gross Domestic Product (GDP) or at least 15-20% of their public expenditure to education, prioritizing groups most in need; and strengthen financial cooperation for education, prioritizing countries most in need.

Source: Author's compilation.

**Table A.2. Proposed targets and goals – UN workstream, 2013-2014**

| Input and source | Date of publication | Proposed education goal or reference to education | Proposed education targets |
|---|---|---|---|
| *Thematic consultation and Technical Support Team Issue Briefs*<br><br>(UNICEF/UNESCO, 2013)<br>(UNGA, 2014a) | September 2012/March 2013<br><br>June 2013 | Equitable Quality Education and Lifelong Learning for All | - All girls and boys are able to access and complete quality pre-primary education of an agreed period (at least 1 year);<br>- Equal access to and completion of a full course of quality primary schooling, with recognized and measurable learning outcomes, especially in literacy and numeracy.<br>- All adolescent girls and boys access and complete quality lower secondary/secondary education with recognized and measurable learning outcomes;<br>- All youth and adults, particularly girls and women, access post-secondary learning opportunities for developing knowledge and skills, including technical and vocational, that are relevant to the worlds of work and life and necessary for further learning and forging more just, peaceful, tolerant and inclusive societies |
| *Co-Chairs' Summary bullet points from OWG-4*<br><br>(OWG, 2013a; see also OWG, 2013b) | June 2013 | Complete the unfinished work of the MDGs to ensure universal primary school enrollment - but also learning outcomes, relevance to job needs, lifelong learning, adult literacy, and non-formal education. | |
| *Co-Chair's summary on meetings with the Major Groups and* | June 2013 | Participants highlighted that education is at the core of sustainable development and the most effective way to poverty eradication. Education as an enabler for progress in other fields was stressed, including health and employment. Need to shift focus from the mere access to education to quality education, including adequate facilities, qualified teachers, good home-conditions, promotion of innovation and | |

| | | |
|---|---|---|
| *other Stakeholders*<br>*OWG-4*<br>(OWG, 2013c) | | civic- mindedness as well as measurable learning outcomes, was stressed by many. So was the need for free and equal access to education for all, including people with disabilities, children living in remote rural areas, and marginalized groups. Importance of life long learning was also raised. |
| *Focus Area Document*<br><br>(OWG, 2014a) | February 2014 | Everyone has a right to education, which opens up lifelong opportunities and is critical to achieving poverty eradication across generations. Achieving universal access to and quality of education is also important in promoting gender equality and empowerment of women, and in shaping values and creating the necessary skilled and productive labour force. Some areas that could be considered include:<br>- Universal primary education for girls and boys, significant progress towards ensuring that every child receives at least a secondary education, and lifelong learning opportunities;<br>- Ensuring equitable access to education at all levels with focus on the most marginalized; achieving high completion rates at all levels of education;<br>- Ensuring effective learning outcomes at all levels and imparting knowledge and skills that match the demands of the labour market, including through vocational training;<br>- Universal adult literacy; improving access to education for persons with disabilities;<br>- Extending where needed opportunities for early childhood education; and<br>- Integrating sustainable development in education curricula, including awareness raising on how diverse cultures advance sustainable development. |
| *Compendium of existing goals and targets under the 19 Focus Areas being considered by the Open Working Group*<br><br>(OWG, 2014b; see also OWG 2014c) | March 2014 | - Achieve high completion rates at all levels of education<br>- Ensure equitable access to education at all levels with focus on the most marginalized<br>- Ensuring effective learning outcomes at all levels and imparting knowledge and skills that match the demands of the labour market, including through vocational training<br>- Ensuring equitable access to education at all levels with focus on the most marginalized<br>- Extending where needed opportunities for early childhood education<br>- Improving access to education for persons with disabilities<br>- Integrating sustainable development in education curricula, including awareness raising on how diverse cultures advance sustainable development<br>- Universal adult literacy<br>- Universal primary education for girls and boys, significant progress towards ensuring that every child receives at least a secondary education, and lifelong learning opportunities |

Source: Author's compilation.

**Table A.3. Proposed targets and goals – OWG and Intergovernmental negotiations, 2014-2015**

| Document and OWG session | Overarching goal | Targets |
|---|---|---|
| *Working Document for the 11th session of the Open Working Group* (May 2014)<br><br>(OWG, 2014d). | Provide quality education and life-long learning for all | a. By 2030 ensure universal, free, equitable access to and completion of quality primary and secondary education for all girls and boys, leading to effective learning outcomes.<br>b. Ensure that persons with disabilities have access to inclusive education, skills development and vocational training.<br>c. By 2030 increase by x% the proportion of children able to access and complete quality pre-primary education.<br>d. By 2030 achieve universal youth and adult literacy, with particular attention to women and the most marginalized.<br>e. By 2030 increase by x% the number of young and adult women and men with vocational training, technical, engineering and scientific skills.<br>f. Integrate relevant knowledge and skills in education curricula, including ICT skills30, education for sustainable development, and awareness raising on culture's contribution to sustainable development.<br>g. All schools to provide safe and healthy learning environment for all students.<br>Appropriate means of implementation. |
| *Proposed Goals and Targets - 12th session of the Open Working Group* (June 2014)<br><br>(OWG, 2014e) | Provide equitable and inclusive quality education and life-long learning opportunities for all | 4.1. By 2030 ensure all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes.<br>4.2. By 2030 ensure equal access for all to affordable quality tertiary education and life-long learning.<br>4.3. By 2030 increase by x% the proportion of children able to access and complete inclusive quality pre-primary education and other early childhood development programmes.<br>4.4. Achieve universal youth literacy and basic numeracy and an increase by x% of adult literacy and basic numeracy by 2030.<br>4.5. By 2030 increase by x% the number of young and adult women and men with the skills needed for employment, including vocational training, ICT, technical, engineering and scientific skills. |

| | | 4.6. | By2030 ensure that people in vulnerable situations and marginalized people including persons with disabilities and indigenous peoples have access to inclusive education, skills development and vocational training aligned with labour market needs. |
| --- | --- | --- | --- |
| | | 4.7. | By 2030 integrate relevant knowledge and skills in education curricula and training programs, including education for sustainable development and awareness raising on culture's contribution to sustainable development. |
| | | 4.8. | By 2030 ensure that all schools and other educational institutions provide safe, healthy, non-discriminatory and inclusive learning environments for all. |
| | | 4.9. | By 2030 enhance the quality of teaching through promoting training for teachers. |
| *Proposed Goals and Targets - 13th session of the Open Working Group* (July 2014)<br><br>(OWG, 2014f) | Ensure inclusive and equitable quality education and promote life-long learning opportunities for all | 4.1. | By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes. |
| | | 4.2. | By 2030 ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education. |
| | | 4.3. | By 2030 ensure equal access for all women and men to affordable quality technical, vocational and tertiary education, including university. |
| | | 4.4. | By 2030, increase by x% the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship. |
| | | 4.5. | By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples, and children in vulnerable situations. |
| | | 4.6. | By 2030 ensure that all youth and at least x% of adults, both men and women, achieve literacy and numeracy. |
| | | 4.7. | By 2030 ensure all learners acquire knowledge and skills needed to promote sustainable development, including among others through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and non-violence, global citizenship, and appreciation of cultural diversity and of culture's contribution to sustainable development. |

| | | 4.a. Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all. |
| --- | --- | --- |
| | | 4.b. By 2020 expand by x% globally the number of scholarships for developing countries in particular LDCs, SIDS and African countries to enrol in higher education, including vocational training, ICT, technical, engineering and scientific programmes in developed countries and other developing countries. |
| | | 4.c. By 2030 increase by x% the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially LDCs and SIDS. |
| *Transforming our world: the 2030 Agenda for Sustainable Development. Resolution adopted by the General Assembly* (September 2015)  (UNGA, 2015) | Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all | 4.1. By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes.  4.2. By 2030, ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education.  4.3. By 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university.  4.4. By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship.  4.5. By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations.  4.6. By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy.  4.7. By 2030, ensure that all learners acquire the knowledge and skills needed to promote sustainable development, including, among others, through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and non-violence, global citizenship and appreciation of cultural diversity and of culture's contribution to sustainable development.  4.a Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all. |

|  | 4.b By 2020, substantially expand globally the number of scholarships available to developing countries, in particular least developed countries, small island developing States and African countries, for enrolment in higher education, including vocational training and information and communications technology, technical, engineering and scientific programmes, in developed countries and other developing countries. |
|  | 4.c  By 2030, substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing States. |

Source: Author's compilation.