W

The construction of SDG4: Reconciling democratic imperatives and technocratic expertise in the making of global education data?

Clara Fontdevila

ORCID: 0000-0003-0589-558X and

Sotiria Grek

ORCID: 0000-0001-5452-7762

The construction of SDG4: Reconciling democratic imperatives and technocratic expertise in the making of global education data?

Abstract

One of the most recent and ambitious exercises in global education monitoring is the construction of the Education 2030 agenda, or Sustainable Development Goal 4 (SDG4). The making of the SDG4 represents the single biggest attempt to reach a consensus on a number of universal education indicators, as well as on the appropriate methodologies and data sources. It is a country-led process, steered by UNESCO but with the input of all major international organizations (IOs). Given its collaborative nature, it confronts its participants with technical but also political challenges.

Through an in-depth analysis of texts and interviews, this chapter enquires into the tension between technocratic legitimacy and political legitimation, by grappling with the question of how expertise and trust in objectivity can be maintained in political processes that aspire to greater inclusion. Particular attention is paid to UNESCO as the custodian agency of SDG4 with a double accountability obligation to participating countries: first, the robust and objective monitoring of progress towards SDG4; and secondly, the participatory and democratic character of such a process. As the chapter shows, this double responsibility does not happen without friction – it has indeed caused tension between and within the participating IOs, as well as between IOs and other stakeholders.

Introduction

Datafication in education governance has become established in recent decades as the prime mode of knowing and reforming complex education systems around the world. The rise of large international assessments created a wealth of statistical information and thus allowed states and transnational agencies for the first time to construct comparative knowledge about

W

education performance. One of these global education monitoring exercises is the construction of the Education 2030 agenda, otherwise the Sustainable Development Goal 4 (SDG4), which promises to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (United Nations General Assembly, 2015, p. 14). SDG4 represents the single biggest attempt to bring together a vast array of actors and countries in order to construct universal education indicators, as well as to decide on the appropriate methodologies and data sources. It is a country-led, global exercise – led by UNESCO, but with the collaboration and close involvement of all major international organizations (IOs). Given its global and collaborative scope, it presents its leaders and participants with enormous technical and political problems.

Through an in-depth analysis of texts and interviews, this chapter will discuss the conundrum of securing accountability of this global performance monitoring project through ensuring the objective validity of its measurement tools, whilst promoting the democratic and equal participation of all actors. UNESCO, as the custodian agency of SDG4, has a double accountability obligation to participating countries: first, the robust and objective monitoring of progress towards the SDG4 goals; and secondly, the participatory and democratic, equitable process in which all member countries have a voice and stake in the project. As a result, although the UNESCO Institute of Statistics has been significantly reinvigorated in relation to its statistical capacity, it has also put great emphasis on the participatory, inclusive and consensual aspects of the agenda – a process that some of the participants see sometimes as taking place at the expense of the robustness of the data produced. As this chapter will show, this double responsibility does not always happen without friction; on the contrary, it has been causing a significant amount of tension in the relationship between and within some of the major IOs, as well as between IOs and other stakeholders, including developing nations.

One of the key questions that this chapter aims to answer is where the accountability of these large performance monitoring projects lies. As education governance by data continues expanding from cradle to grave, can governing through data secure both technocratic legitimacy and political legitimation at the same time? Does the democratic expansion of participating actors pose threats to technocracy as a scientific, robust endeavour? This chapter grapples with the question of how expertise and trust in objectivity can be maintained in political processes that aspire to greater inclusion and the set-up of common, global objectives. How do the processes of indicator development change as they

W

are increasingly subject to democracy and transparency demands? Which compromises are needed in order to reconcile international and national data sources relying on different sources of expertise? Ultimately, how can we theorize on the relationship of elite, expert, technocratic objectivity-making with the inclusive, participatory and ever-expanding process of making consensus?

This chapter will begin with a short overview of the main theoretical underpinnings of the notions of accountability and transparency and their relationship with the growing role of datafication in education. We will continue with a focus on the historical trajectory of the initial inception, followed by the first steps and later the decisive moments that made the SDG4 the kind of monitoring instrument that it is today. Next, we will offer a short description of the research design and methods, before moving on to the analysis of interview data.[1] The chapter will show how the notions of mediation and instrumentation, as they emerge from actors' accounts, point towards a better understanding of the uneasy coexistence of different accountabilities in the transnational governance of education.

Accountability and transparency in the age of governing by numbers

Since the 1990s, both transparency and accountability have become dominant tools in the education governing arsenal; indeed, Meyer et al. (1997) see them having been imposed as a new world society norm. Especially in the field of the transnational governance of education, rituals of comparison and performance measurement as well as the economization of education discourses have become integral modes of the universal panacea of governing by a distance (Ozga, 2013). New forms of accountability and audit have transformed national education governing regimes, too, and in doing so, they have directly impacted on the transnational architecture of education.

Indeed, Power has insightfully written about the rise of the audit society (1999) and the enabling of calculative technologies that allow for decision-making and action at a

[1] This publication is informed by two separate research projects: (a) it is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 715125 METRO (ERC-2016-StG) ('International Organisations and the Rise of a Global Metrological Field', 2017–2022, PI: Sotiria Grek); and (b) it draws on the data collected in the context of a PhD project funded by the Spanish Ministry of Education and Science under the Programme for University Lecturer Formation (FPU) (grant FPU014/0611, awarded to Clara Fontdevila).

W

distance (Porter, 1996). Although these accountability structures are power-laden and hierarchical, these hierarchies are not rigid or stable (Djelic & Sahlin-Andersson, 2006). The endless "treadmill of accountability" (Bostrom & Garsten 2008, p. 242) requires that actors and alliances are permanently in movement and are always incomplete. Thus, accountability can arguably by equated to a normative shell through which audit, standards and deregulation are pushed in education.

Given the increasing pervasiveness and centrality of accountability and datafication in education governance, a number of theoretical tools have been developed to understand such transformations. A recent contribution to this scholarship – that proves particularly apposite to the objectives of this chapter – is Maroy and Pons' (2019) tripartite schema for the examination of the *practice* of accountability. The relevance of this work lies in the fact that it goes beyond the performativity perspective that has dominated education scholarship from the 1990s onwards. Through the application of the theoretical tools of the *sociologie de l'action publique*, Maroy and Pons invite us to examine not only the accountability policies *per se*, but in fact pay attention to the notions of *trajectory*, *mediation* and *instrumentation* in school accountability:

- The notion of *trajectory* is used to make sense of the emergence and implementation of a given policy instrument. According to Maroy and Pons (2019), this needs to be understood in relation to three intertwined processes and factors: namely, an element of path-dependency; bricolage practices through which available tools are recombined; and the translation and recontextualization of tools and ideas developed elsewhere and, particularly, those circulating at a transnational level.
- The concept of *mediation*, in turn, captures process of transformation and mutation. This is a consequence of the co-construction work performed by actors operating at different levels and in different settings. The notion throws light on the interplay and the combined impact of a wide range of organizations and individuals – whose action contributes to (re)define the meaning of policy tools.
- Finally, the process of *instrumentation* borrows from political science debates on the making and role of policy instruments, i.e. "an apparatus that is both technical and social, that organizes specific social relations between the state and those it is addressed to, according to the representations and meanings it carries" (Lascoumes & Le Galès, 2004, p. 13). It captures thus the dynamics through which policy instruments come into being.

W

Before moving on to a short description of our methodology and the examination of actors' accounts, the next section will focus on a description of the historical context of the emergence of the SDG4. Through exploring the *trajectory* of initially an idea devised by a small set of people to the major monitoring instrument that it has now become, the following account shows the degree to which a certain path dependency, coupled by some bricolage work, gave birth to the core conundrum surrounding the SDG4: technical or democratic?

Negotiating and producing metrics for Sustainable Development Goal 4: A historical trajectory

2015 was a crucial year for the global education community. In May, a World Education Forum (WEF) was celebrated in Incheon (Republic of Korea) with the participation of over 1,500 participants, including 120 Ministries of Education and representatives from a wide range of international governmental and non-governmental organizations. The gathering was devised (and touted) as a successor or the Jomtien and Dakar meetings celebrated respectively in 1990 and 2000 – both of them widely acknowledged as milestones in the development and consolidation of the Education for All movement. The main product of WEF 2015 was the so-called Incheon Declaration, along with the Framework for Action adopted by UNESCO Member States in November 2015. In conjunction, both documents established an ambitious and highly aspirational education agenda for the 2015–2030 period, condensed in the overarching goal "Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" and a number of associated targets (UNESCO, 2016).

What is important here is that both documents were the product of a long, multilayered and multisited negotiation process that involved numerous meetings and consultations, largely led by UNESCO under the auspices of the EFA Steering Committee. At the same time, it should be noted that efforts towards the development of this agenda were in turn paralleled by the negotiation of the Sustainable Development Goals – one of the cornerstones of the 2030 Agenda adopted by the General Assembly of the United Nations in September 2015 and devised as a follow-up of the Millennium Development Goals. In fact, the EFA-led process and the debates facilitated by the UN Open Working Group on Sustainable Development Goals reinforced and informed one another through an intricate political process, eventually crystallizing in a single agenda conventionally known as SDG4/Education 2030 – a denomination reminiscent of the dual origins of the new set of goals.

W

Education 2030/SDG4 represented simultaneously a form of continuity and a departure from previous instances of goal-setting such as EFA and the Millennium Development Goals. As a programmatic document oriented at nurturing and securing a form of collective commitment towards a shared set of aspirations, the new agenda builds on a well-established tradition that has come to be recognized as a procedural hallmark of the UN system. However, Education 2030 entails a certain discontinuity regarding education goal-setting practices – in both content and procedural terms.

First, and as different scholars have noted, the new set of goals is characterized by an unprecedented degree of ambition – shifting away from the focus on primary education and gender equality that characterized the MDG era, but also expanding on the vision set up by the EFA program; and establishing a truly universal agenda that contrasts with the prior focus on developing countries (King, 2017; Unterhalter, 2019). Secondly, the very *making* of Education 2030 (and of the SDGs more in general) represents a path-breaking development in the long history of goal-setting practices and UN summitry. The open, inclusive and participatory nature of the consultative process facilitated by UNESCO and the EFA architecture was in many ways unprecedented, and the openly-negotiated and improvisatory character of the SDG debate contrasted with the technocratic origins of the MDGs (cf. Fukuda-Parr & McNeill, 2019).

In addition to these normative and procedural shifts, the SDGs agenda introduced also a greater emphasis on accountability and transparency – increasingly portrayed as key operating principles and as governance challenges directly influencing the implementation of the SDG framework (Bowen et al., 2017). Such commitment contrasts with the diffuse lines of responsibility that characterized the MDG era (Clegg, 2015). It is significant, for instance, that the notion of accountability features prominently in the roadmap sketched by the Education 2030 Framework for Action. At the same time, it should be noted that the policy translation of such principle remains largely undefined. The notion of accountability seems to be equated to an emphasis or commitment to transparency, and to efforts to improve monitoring, reporting and follow-up practices. The focus on shared responsibility makes it difficult to ascertain who should be accountable to whom, and over which outcome. In practice, the implementation of the accountability principle in relation to the SDG agenda has materialized in the adoption of a series of monitoring and review mechanisms – most notably, the delivery of formalized and periodic reviews at the national, regional and international levels (Bowen et al., 2017). In this new scenario, data-collection

W

efforts have only been gaining prominence – with two of the targets of SDG17 explicitly aiming to affect and improve the availability of data, to develop new measurements of progress on sustainable development, and to strengthen the statistical capacity of developing countries.

Although, as noted above, 2015 was much of a turning point for the global education community, it left a number of issues open – among them, the so-called *indicator debate.* The fact that the measurement of SDG4 remained an instance of "unfinished business" is all more relevant given the centrality placed on data-collection, monitoring and tracking efforts. Over the last years, the datafication of transnational education governance has come to be perceived as the sole means through which the so-called "learning crisis" can be solved. Brought to the fore by the 2013/2014 EFA Global Education Monitoring, the idea of a global learning crisis aims at capturing and exposing the limited academic progress made by schooled children in a number of low- and middle-income countries. As noted by Barrett (2016), the emergence of this notion was in fact made possible by the proliferation of learning assessment. The author goes on to argue that the so-called testing culture and the learning crisis narrative have tended to reinforce each other – for solving the crisis requires data, and such data only renders more evident the extent of the crisis.

Reflecting the primacy given to data-collection and monitoring efforts, the Framework for Action established four levels of indicators (global, thematic, regional and national) associated to different reporting strategies. In the light of the unequal development of these indicators, and their availability and coverage at a global scale, a tier classification tool was implemented. Importantly, a number of custodian agencies became responsible for the development and refinements of such indicators. In the case of education, the UNESCO Institute for Statistics (UIS) became responsible for 9 out of 11 indicators – sharing the responsibility with UNICEF and the OECD for the other remaining two. Given the initial classification of a number of metrics as Tier 2 and Tier 3 indicators (e.g. indicators not regularly produced by countries or for which measurement standards are not available yet), their refinement and production rapidly become a priority for UIS, which, as discussed below, perceived their organizational legitimacy or prestige as closely tied to the fate of these global indicators (cf. UIS, 2017a).

Global indicators were in turn complemented by a thematic indicator framework, originally developed by the Technical Advisory Group on Education Indicators (TAG) established by UNESCO in 2014 and included in a draft form in the Framework for Action.

W

This document proposed up to 43 indicators and, more importantly, established the leading role of the UIS in developing and refining this initial list – in collaboration with Member States and other education stakeholders and SDG4 partners.

Given the complexity of such endeavours, but also in order to guarantee the participation of a wide range of stakeholders, two ad hoc mechanisms were created with a view to advance the development and production of SDG4 global and thematic indicators – namely, the Technical Cooperation Group (TCG) and the Global Alliance to Monitor Learning (GAML). The former was established in 2016, being conceived as a space for discussion as well as a technical platform to support UIS in the implementation of the thematic indicator framework, and to assist other 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 by regionally-representative UNESCO Member States, as well as representatives of different IOs, civil society organizations and the co-chair of the Education 2030 Steering Committee (see UIS-TCG, 2017).

The GAML, in turn, was created in 2016, being originally defined as an "umbrella initiative to monitor and track progress towards all learning-related Education 2030 targets" (UIS, 2016, p. 49), and tasked with the development of tools, methodologies and shared standards to measure learning outcomes in the context of SDG4. The GAML is also expected to support and strengthen country-level capacities to generate and use learning data (see UIS, 2017b). Its membership is open to any individual or organization willing to contribute to the work of the GAML. Representatives of Member States have been gaining prominence since 2018, with the engagement of countries being increasingly perceived as central to the GAML's success and legitimacy. With the UIS hosting the Secretariat, the GAML operates by definition in an open and participatory manner, with decisions being made through consensus.

Importantly, these mechanisms did not emerge in a vacuum. Quite the contrary, and as the notion of policy trajectory helps establish, both of them were born out of already existing initiatives launched during the run-up towards the approval of Education 2030. More specifically, the TCG represents a continuation of the TAG established in 2014. The GAML, in turn, was a successor of the Learning Metrics Task Force (LMTF), launched in 2012 as a multi-stakeholder partnership co-convened by the Center for Universal Education at the Brookings Institution and the UIS. However, both the TCG and the GAML entailed a

W

procedural shift vis-à-vis their own precedents. Both initiatives are explicitly articulated as subject to a transparency mandate, and are expected to operate through democratic, equitable and inclusive dynamics. In this sense, both spaces are subject to a dual form of accountability; they are held responsible over the results of their work, but judged also over the quality and inclusive character of their deliberations. To put it differently, both spaces are characterized by a built-in tension between a democratic imperative and the technical and specialized nature of the work they are tasked with.

In the following section, we examine the manifestations and implications of this tension between a democratic and a technical ideal. To this end, we draw on data from semi-structured interviews with representatives of different organizations partaking in the SDG4 indicator development efforts – including IOs, but also civil society organizations and technical partners actively involved in such endeavour. The interviews were conducted during the 2017–2019 period in the context of two separate, but complementary studies.[2] Such data-collection efforts were in turn complemented with the documentary analysis of publicly-accessible reports, technical notes and declarations produced in the context of the SDG4 negotiations and, specifically, indicator-development efforts as those that centred the activity of the GAML and the TCG.

From mediation to instrumentation: Producing data and consensus for the SDG4

As the overview above suggests, the negotiation and production of SDG4 indicators has been guided from the outset by a double imperative – preserving technical sophistication while ensuring political legitimacy. Even if (theoretically) such principles are not mutually exclusive, as we will show below, this double imperative has frequently translated into a form of tension or duality. Thus, the everyday practice of turning these principles into actionable strategies and plans is often associated with different courses of action, different (and sometimes competing) forms of expertise and different valuation standards and consensus-making practices. Making decisions often means prioritizing among these competing logics – deciding on a division of labour and an agenda, but also determining who is brought to the table, whose proposals are selected and whose proposals are ruled out. It is in these *mediation* practices that the tension between technocratic and political legitimacy

2 Please see note 1.

W

crystallizes and becomes evident. In this section, we inquire into the different manifestations of such mediations and translations in bridging concerns of the apparent contradiction of a highly technical process, being also a democratic one.

Whose voices? Opportunities and costs of "getting everyone on board"

Debates around the role of country representatives are particularly illustrative of such tensions and mediations. The inclusion of country representatives to the Technical Advisory Group,[3] and the transition from the TAG to a TCG explicitly geared towards the incorporation of national voices, constitute revelatory episodes of the frictions created by this democratization ideal. The unease generated by the incorporation of country representatives among UN agencies appears to stem partially from the (perceived) risk of politicization of what would otherwise constitute a dispassionate, technical debate. An additional source of concern is the nature of the expertise exhibited by those representing the national perspective. Country representatives are thus perceived as lacking the necessary preparation to effectively contribute to TCG debates:

> There's of course the other issue of people who come tend to be illiterate statistically. Not always, because some countries send bureaucrats. But there are enough people. But the question is how are these people linked, and how well do they represent first of all not just their own country, but the region (UNESCO 1)

The evolution of GAML suggests similar dynamics. Widening the membership of the platform (and especially incorporating country representatives) was seen as necessary to secure the buy-in and cooperation on the part of the Member States that, ultimately, should produce and developed the indicators, and report them according to the procedures and standards discussed in GAML:

> The first meeting of GAML only had technical partners around the table. And I would say to an extent this is 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. (Interviewee/UIS)

[3] Originally formed exclusively from representatives from IOs and technical agencies.

W

At the same time, such democratization efforts were perceived as problematic by those participants with a more technical profile – or more concerned by the statistical and psychometric sophistication of data-collection efforts:

> We can't consult everyone, everywhere, we've gone through this whole technical consultation. Is there some value in saying these are the technical standards? … And it keeps going back to this theme of democracy and voice. Is there such a thing as too much democracy, too much consultation? At what stage do you say we just have to run with this, this is what it is? (Interviewee/World Bank)

Paradoxically, what was perceived as *excessive* consultation among certain GAML members, was deemed insufficient among others – especially those representing civil society more explicitly driven by a principled agenda privileging openness and country ownership. The inclusion of these domestic actors was sometimes perceived as a performative exercise motivated by formalist concerns. Given that the preoccupations or motions put forward by countries do effectively inform the work or decisions made by GAML, some interviewees pointed to a risk of tokenism:

> And then what the GAML did last time was that they actually supported representatives from the Global South, but they were flown in to give presentations about what they are already doing. Which isn't a bad thing to do, but in practice it meant that we had about six hours of uninterrupted PowerPoint presentations and no real participation of these people in the actual meeting. … It was a bizarre setup. (Interviewee/Civil society)

Whose data?

When it comes to the production of indicators, the selection or prioritization of specific data sources has also brought to the fore the tension between democratic and technocratic ideals. While both of them are expected to equally inform the measurement labour of UIS, the selection of one dataset or another frequently entails a de facto prioritization of one principle over the other. Disputes on the measurement of learning through large-scale assessments are particularly illustrative of such tension. While some IOs and assessment consortia claim that the need for global reporting in an argument in favour of the in-built comparability enjoyed by *international* large-scale assessments (ILSAs), other actors have called into question such notions – arguing that the value of data lies in its *usability for domestic policymakers* and that, in consequence, national large-scale assessments (NLSAs) represent equally valid (if not more appropriate) data sources when it comes to measuring learning. Ultimately, the

W

ILSAs vs. NLSAs debate echoes a more fundamental disagreement on what constitutes valuable or trustworthy knowledge:

> We [*at the OECD*] are working with UNESCO, and the UIS, to help fill capacity for measuring progress at the end of primary. And there's a proposal to use national assessment systems, as far as possible, but a lot of these national assessment systems … the quality of these assessments is not good (Interviewee/OECD)
>
> 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? […] 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 PIRLSS, or you know, one of the international studies (Interviewee/ACER)

Negotiating organizational legitimacy

One of the direct consequences of the existence of two (potentially conflict) principles expected to guide data production has been the emergence of competing sources of organizational legitimacy or authority. Since such forms of authority are enjoyed to differing degrees by the concerned IOs, disagreements over the hierarchization of democratic vs. technocratic demands should be read in connection with the competition for dominant positions in the measurement field. As a consequence, a number of inter-organizational tensions have unfolded – with some organizations privileging technical robustness and others putting a premium on procedural values such as inclusion, democracy and consensus. Among the former, the OECD features prominently. Representatives of the organization have thus argued that the technical expertise of the organization is unmatched in the education measurement realm. Conversely, representatives of UNESCO emphasize their own mandate as the custodian agency for SDG4, as well as their capacity (and responsibility) to represent the interests of all Member States:

> The OECD really doesn't have a mandate to work with Sub Saharan Africa. This is the mandate of the UNESCO institute of statistics. So OECD shouldn't come now and tell even UNESCO for example, you were talking about the relationship between OECD and UNESCO. OECD shouldn't go to UNESCO in Paris and say oh we can do all these things for you, why don't you work with us? There is an Institute of Statistics that's part of UNESCO that should be doing this. (Interviewee/UIS)

W

Importantly, the technocratic-political tension manifests itself in the form of friction or dissent not only *between* organizations, but also *within* organizations. Thus, technocratic and democratic imperatives are likely to translate into conflicting demands that pull certain organizations toward opposite directions. Such dynamics have been particularly visible in the case of the UIS – which, as noted above, is held accountable over the technical robustness of the data produced and reported, but also over the democratic and inclusive character of data-collection and data-production processes. This frequently translates into the need for some form of compromise. It is precisely in this form of compromise (both internally and externally) that UIS as an actor has had an important mediating role; with UNESCO being the lead organization, the independence of UIS puts is in a position where it needs to mediate between both the technical experts of other IOs, in addition to the numerous public and private research agencies that contribute with data, *and* national representatives and civil society actors who are interested in the democratic legitimation and equal representation of all countries involved. Although brokerage efforts for finding some kind of equilibrium between technical superiority and plurality of voice are to be found in all IOs involved, the UIS – with its proclaimed independence from UNESCO – is the key institution mediating and acting as the buffer in this unusual epistemic vs. political tug-of-war. In the following interview extracts, UIS actors discuss the opportunities but also difficulties of having this role:

> It has always been very difficult to have a global conversation about these tools that are being used to measure learning in different regions […] And the UIS neutrality in the global governance agenda, plus its mandate, positions the UIS as the natural convener for the discussion. And many stakeholders do agree that the UIS is neutral and has this mandate, that gives the UIS the leverage to engage with this discussion. So I think it would have been unfortunate for the UIS to let this opportunity go (Interviewee/UIS)
>
> So OECD has a TCG, but the world has not a TCG of the same calibre, where countries lead the decision. And I'd say it's very, very difficult. To the credit of the UIS for having established it … but the amount of effort it would take to do that at a global level is very difficult, is very high. So the problem is the effort it takes. You know, the UIS is not a particularly well-funded organization … It has a convening power, but it doesn't have the resources (Interviewee/UNESCO).

W

At the same time, the reliance on consensus-making and the participatory, inclusive nature of UIS-led debates could come at a cost. The UIS risk being perceived as inefficient as a result of its own democratization efforts. Such reputation is considered to negatively impact the volume of resources available for measuring and monitoring SDG4:

> My take on that, and I could be wrong, but my take is the funding will be available … and the funding is out there, provided GAML is clear about their governance and what it wants to achieve. So I think that among the funders and government agencies and philanthropic agencies and so on, there's plenty of money to do this. That would be my take. But, these people are gonna keep their hands in their pockets, not pull the cash, you know, until it's not absolutely clear what's gonna be done, by whom it's gonna be done, and where's gonna be done. (Interviewee/Technical partner)

Discussion

After decades of education scholarship examining accountability as fluid, chameleonic and all-pervasive (Sinclair, 1995), SDG4 represents a clear case of a new era in the role and practices of accountability – at least at the level of the transnational governance in education. The chapter has discussed the tensions arising from the coexistence of two, separately demarcated, types of accountability logics. As we have seen above, these are the political and the technical types of accountability, which, at least in the governing of the SDG4, appear to be stirring uneasy discussions amongst the main players in the field. The chapter has shown actors' apprehensiveness about whether they can maintain the legitimacy of their expertise, when faced with the demand for technocracy to be politically accountable, too. On the other hand, those actors who strive for more inclusive and diverse decision-making, do not want to be seen as less rigorous, or as compromising the validity and robustness of the data produced. It has to be stressed that although, in the quotations above, the divergence between the UNESCO versus the OECD and World Bank perspectives seem to frame a discussion of contrasting organizational cultures and ways of working, a deeper look into the controversy shows that organizations do not take monolithic or single views on this, but that there is also a great deal of internal organizational struggle and contest.

How to make sense of this open disagreement over which type of accountability ought to prevail in the process of producing monitoring for SDG4? First, our data purport that the making of the SDG4 offers a particularly opportune moment in which to enter the organizational agendas and practices of the major international organizations, whose work has been defining in many ways the trends in global education governance for at least two

W

decades. Both projects are focused on exploring the workings of producing data for education policy; by entering the laboratory of the making of numbers, one can acquire a much better grasp of the political nature of the decisions in relation to the production of data.

Secondly, and perhaps more importantly, the examination of the making of the SDG4 as the first collaborative endeavour to produce a measurement agenda for global governance is a unique opportunity to examine how the major IOs come together and drift apart in the making of data for policy. In many ways, the politics of data production has so far been part of the black box of organizational processes that, in their majority, remain hidden from public (and researchers') view. The interdependency of IOs in producing monitoring for the SDG agenda allows for a more nuanced understanding of the politics of the making of quantification, since it reveals the knowledge and value conflicts that large IOs have.

However, beyond IOs' interdependency and the challenges that come with it, what is it about the SDG4 that unravels the accountability conundrum and requires that different kinds of accountability are treated equally in the governing of the transnational space of education? For a long time, the accountability that large international comparative assessments faced was one that related solely to their robustness and trustworthiness. Thus, comparative assessments have long been challenged primarily on technical grounds – with much of the debate revolving around solidity and validity (see Rutkowski, 2018 for an overview of such limitations). Despite the emergence of more politically-driven criticisms focusing on the (mis)uses and effects of such comparative exercises, the debate on international assessments has tended to develop in a bifurcated way – with political and technical discussions neatly separated one from another, taking place in different spaces and relying on different parlances (Gorur, 2017). What is it about the SDG4 that has unearthed this conflict of political and technical accountabilities? The chapter, through a careful and detailed examination of the historical trajectory and the developments that led to the birth of the SDG4, in addition to actors' own voices, has shown how it was precisely the close collaboration of different IOs that contributed to a large extent to these tensions.

All three key organizations that take part, OECD, UNESCO and the World Bank, bring their knowledge expertise and networks to the table. With the datafication of transnational education governance as a given, the struggle now moves to a different plane: as we have seen in this chapter, some of these IOs see their function as primarily the

W

guarantors of the legitimacy of the whole global education governing agenda as fair and democratic. Thus, UNESCO, at least, appears to be using its capital as a trustworthy defender of a more humanistic face of education, while the OECD and the World Bank, although protecting their technical expertise, also use their networks and tried-and-tested methods of managing consensus. In this new, global game of the datafication of education governance, the concerns of legitimizing the process move now beyond the mere argument of "trust in numbers". As shown above, the struggle of individual and organizational actors is to maintain the scientific authority and technical objectivity of the exercise, whilst ensuring that this is a politically and democratically fair process – and not another colonial endeavour into building new data empires in the Global South.[4]

This chapter argues that Maroy and Pons' (2019) tripartite schema of the examination of *trajectory*, *mediation* and *instrumentation* represents a promising approach to understand the challenges faced by SDG4 it its attempt to bring the two accountabilities together: the political and the technocratic one. Thus, and as showed in detail above, the historical trajectory of the SDG4 played a key role in how it was shaped from the beginning, as well as where it would move towards to. Whilst maintaining some of the features of all such monitoring instruments, for example, the defining and setting of an architecture of quantitative measures of performance to be achieved by a totemic time point (2030), SDG4's trajectory also saw the introduction of quite a new governing style: an open, inclusive, participatory mode of organizing the work, which has involved (and continues to involve) extensive consultations. It was precisely this specific trajectory that meant that translation and brokering work would become such as an essential part of the IOs' work.

The notion of mediation as advanced by the authors has also been key for the discussion of the continuous co-construction of accountabilities, at different settings and times and by different actors. Remarkably, although all relevant IOs need to do this kind of mediating work both for their internal as well as their external audiences, the UIS appears as the main actor tasked with the role to find a middle ground between the centrifugal forces of

[4] This preoccupation echoes recent debates on the specific impact of datafication on middle- and low-income in the reproduction of power asymmetries between the Global North and the Global South. See, for instance, Arora's (2016) insights into the parallels between ambitious data projects unfolding in the South and colonial regimes of surveillance; or Couldry and Mejías' (2019) discussion on data colonialism and the extractive nature associated to contemporary data-collection practices.

W

statistical robustness and democratic legitimacy. We have shown above the challenges and opportunities of this political work (Lagroye, 1997), as the Institute navigate the risks and perform the (maybe impossible) task of trying to satisfy such a diverse array of stakeholders and interests.

Indeed, in order to understand the ability of IOs to validate their work through being both technically and politically accountable, we need to understand the core condition that makes any accountability process successful. Thus, there is never a single space or time that any decisions of this kind are taken. The informality, multiplicity and dynamism of the epistemic and policy networks that bring the SDG4 into being is built precisely on the ability to simultaneously use different discourses and accountabilities, depending on the context and participant actors. Therefore, although there may be meetings and spaces where the complex technical statistical work of validating data for indicators takes place, there are also parallel meetings where these data are presented to national representatives for their approval. These represent the essential consensus-making and mediating work that is a sine qua non for the approval not only of countries, but crucially of funders and donors too.[5]

Finally, the making of SDG4 is in essence the manifestation of an "apparatus" as Lascoumes and Le Galès (2004) describe – and, consequently, can be productively analysed as an instance of an instrumentation process as conceptualized by Maroy and Pons (2019). Both technical and social, SDG4 could be seen as a prime example of a transnational soft regulatory instrument (in the tradition of soft law, i.e. best practices, expert standards, rankings, ratings, audits, quality assurance and the like). As such, it creates competitive and reputational pressures on those participating – including pressures for countries to set up (or invigorate) their statistical infrastructure and educational management information (EMIS) systems. In this sense, the SDG4 represents a singular exercise in that it weaves together two forms of accountability that have traditionally been perceived as disparate, if not irreconcilable – that is, political and administrative accountability (see Bovens, 2007;

[5] Thus, the UIS receives financial support from a variety of bilateral aid programs and philanthropic organizations – relevant donors include the Bill & Melinda Gates Foundation, the UK Department for International Development and the Norwegian Agency for Development Cooperation. While such organizations expect the UIS to make significant, rapid progress in the measurement of SDG4, they also place a high value on an inclusive and participatory modus operandi – one open to the input of a variety of development partners, as opposed to one tightly controlled by international bureaucracies.

W

Verger and Parecerisa, 2017 for a discussion of different forms of accountability in education). This is so as the possibility for countries to engage in a form of political accountability (by reporting on and answering for their progress against the SDGs) is made conditional on their engagement with administrative forms of accountability (through which public administrations collect and analysed data on performance or behaviour of a range of education stakeholders).

Finally, the chapter has shown how the SDG4 does not merely set broad global measurement agendas, but crucially politically legitimizes the endeavour by claiming that it gives participating actors equal say – putting different countries and IOs on an equal footing. Therefore, it is interventionist in nature, leading effectively to a mutually constitutive relationship of the statistics and the countries they are meant to represent. More importantly, it represents a leap in the practice of transnational soft regulation in education because, although prescriptive, it also appears as transparent, pluralistic, open and developmental. In the end, little will it matter whether the targets will be met or not. In having set up such a complex, yet fluid, governing infrastructure, the SDG4 has been successful already.

Funder acknowledgement

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

References

Arora, P. (2016). Bottom of the data pyramid: Big data and the global south. *International Journal of Communication, 10*(2016), 1681–1699

Barrett, A. M. (2016). Measuring learning outcomes and Education for Sustainable Development: The new education development goal. In W. C. Smith (Ed.), *The Global Testing Culture: Shaping Education Policy, Perceptions, and Practice* (pp. 101–114). Oxford, UK: Symposium Books.

Bostrom, M., & Garsten, C. (2008). *Organising transnational accountability*. Chentelham, UK: Edward Elgar Publishing.

W

Bovens, M. (2007). Analysing and Assessing Accountability: A Conceptual Framework. In *European Law Journal*, *13*(4), 447–468.

Bowen, K. J., Cradock-Henry, N. A., Koch, F., Patterson, J., Häyhä, T., Vogt, J., & Barbi, F. (2017). Implementing the "Sustainable Development Goals": Towards addressing three key governance challenges - collective action, trade-offs, and accountability. *Current opinion in environmental sustainability*, *26-27*, 90–96.

Clegg, L. (2015). Benchmarking and blame games: Exploring the contestation of the Millennium Development Goals. *Review of International Studies*, *41*(5), 947–967.

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, *20*(4), 336–349.

Djelic, M. L., & Sahlin-Andersson, K. (2006). *Transnational Governance: Institutional Dynamics of Regulation*. Cambridge, UK: Cambridge University Press.

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

Gorur, R. (2017). Towards productive critique of large-scale comparisons in education. *Critical Studies in Education*, *58*(3), 341–355.

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

Lagroye J. (1997) *Sociologie politique*, Paris: Dalloz-Presses de la FNSP

Lascoumes, P., & Le Galès, P. (2004). *Gouverner par les instruments*. Paris: Presses de la Fondation Nationale des Sciences Politiques.

Maroy C. & Pons, X. (2019) Introduction. In C. Maroy & X. Pons (Eds.), *Accountability Policies in Education: A Comparative and Multilevel Analysis in France and Quebec*, (pp. 1–12). Cham, Switzerland: Springer.

Meyer, J. W., Boli, J., Thomas, G. M., & Ramirez, F. O. (1997). World society and the nation-state. *American Journal of Sociology*, *103*(1), 144–181.

Ozga, J. (2013). Accountability as a policy technology: Accounting for education performance in Europe. *International Review of Administrative Sciences*, 79(2), 292–309.

Porter, T. M. (1996). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.

Power, M. (1997). *The Audit Society*. Oxford, UK: Oxford University Press.

W

Rose, N., & Miller, P. (1992). Political power beyond the state: Problematics of government. *British journal of sociology*, *43*(2), 173–205.

Rutkowski, D. (2018). Improving international assessment through evaluation. *Assessment in Education: Principles, Policy & Practice*, *25*(1), 127–136.

Sinclair, A. (1995). The chameleon of accountability: Forms and discourses. *Accounting, Organisations and Society*, *20*(2/3), 219–237

UNESCO. (2016). *Education 2030: Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4*. (ED-2016/WS/28). Retrieved from: https://unesdoc.unesco.org/ark:/48223/pf0000245656

UNESCO Institute for Statistics (UIS). (2016). *Sustainable Development Data Digest. Laying the Foundation to Measure Sustainable Development Goal 4.* Montreal, Canada: UNESCO-UIS.

UNESCO Institute for Statistics (UIS). (2017a). *SDG 4 Monitoring: Framework development* (GAML4/REF/18). Retrieved from: http://uis.unesco.org/sites/default/files/documents/gaml4-sdg4-monitoring-framework-development.pdf

UNESCO Institute for Statistics (UIS). (2017b). *Global Alliance to Monitor Learning (GAML): Concept Paper*. Retrieved from: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/10/gaml-concept_paper-2017-en2_0.pdf

UNESCO Institute for Statistics (UIS)/ Technical Cooperation Group (TCG). (2017). *Technical Cooperation Group on the SDG 4 - Education 2030 indicators (TCG). Terms of Reference - Draft.* Retrieved from: http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2018/08/TCG_ToRdraft_20170929.pdf

United Nations General Assembly (2015). *Transforming our World: The 2030 Agenda for Sustainable Development.* Resolution A/RES/70/1. Retrieved from: http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

Unterhalter, E. (2019). The many meanings of quality education: Politics of targets and indicators in SDG 4. *Global Policy*, *10*, Suppl. 1, 39–51.

Verger, A., & Parecerisa, L. (2017). *Accountability and education in the post-2015 scenario: International trends, enactment dynamics and socio-educational effects. Background paper prepared for the 2017/8 Global Education Monitoring Report*. Retrieved from: https://unesdoc.unesco.org/ark:/48223/pf0000259559