PISA and PISA for 'development': an inquiry into the OECD's expansion into low- and middle-income countries

By Xiaomin Li

Thesis submitted to University College London (UCL)

Institute of Education (IOE)

For the degree of

Doctor of Philosophy

January 2021

Declaration

I, Xiaomin Li, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis analyses PISA-D to understand its origins, the legitimation strategies which frame it, and its outcomes. The thesis begins by locating PISA-D within the evolution of OECD after the 1950s. That story, which is told briefly, is one of successful expansion. More specifically, as the global education agenda was being changed after 2015 to emphasise minimum standards of quality for all countries to be delivered by 2030, the OECD has been seeking to expand its most successful policy instrument – PISA – to include low- and middle-income countries. In 2013, it introduced PISA for Development (PISA-D) to help promote PISA as a universal measure of learning. In 2018 the results of PISA-D were released, and in 2019 it was declared a success by OECD.

Clearly PISA-D has the potential to have a significant impact on various aspects of education in low-income countries. To explore these themes, this thesis was guided by a number of research questions which included: what are the origins of PISA-D and how have they shaped its contemporary form? How has PISA-D been promoted and what strategies have been employed by OECD to establish a clear role for itself in monitoring Sustainable Development Goal 4? How has PISA-D addressed the challenges that low- and middle-income countries face in engaging in PISA and what policy insights does it provide for the piloting countries?

The basic approach of the thesis is through documentary research, but this was supplemented by semi-structured interviews and secondary analysis of the literature. The work has permitted new interpretations of the contributions of OECD to global educational governance, has highlighted the legitimation strategies of OECD as it expands its role, and has made clearer how the PISA-D countries are construed as backward – and as needing to learn from other, especially high-performing, countries participating in PISA.

Impact Statement

This thesis focuses on the global expansion of an organisation that is having a significant impact on education policy and research. While there is extensive literature on the role and mechanisms of the Organisation for Economic Co-operation and Development (OECD) in global educational governance, few researchers have analysed the recent PISA for Development (PISA-D) project, which aims to extend the OECD's power and influence into low- and middle-income countries. This thesis addresses that need and provides a critical understanding of the origins of PISA-D, the strategies employed by the OECD to promote and legitimise the project, and the project's outcomes.

The analyses presented in this thesis hold the potential for impact within academia by contributing to knowledge in the fields of comparative education (CE), education and international development (EID), international relations (IR), and organisational studies. My analyses contribute to CE and EID literature with a particular focus on the extension of an 'applied' form of comparative education, as represented by PISA, to developing countries. This is portrayed by the OECD as a development project that contributes to both educational and economic outcomes. My findings concerning the origins of PISA-D and the legitimation strategies of OECD offer critical insights into the changing role played by this organisation and its approach to education and development in the field of global educational governance.

The potential impact of this study extends far beyond academia, since the results of PISA-D will be used as an important source of data for monitoring progress towards Sustainable Development Goal 4—a commitment by the United Nations and its member states to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all by 2030. The results will further provide participating countries with evidence of what works to guide their education reforms and the allocation of domestic and external resources. The OECD is not the sole actor. Other development agencies, such as the World Bank and UNESCO, have similarly adopted this approach of gauging student performance on large-scale assessments (LSAs) as a lever of development. My examination of the outcomes of PISA-D helps in determining whether this specific approach can bring about change in under-resourced settings.

On a personal level, this study has had an impact on my own intellectual and professional development. In particular, this research experience has taught me how to

think critically by acknowledging the multiple perspectives on reality and, therefore, not simply accepting what has been normalised. This ability has enabled me to perceive and analyse the work of OECD and of other prominent development agencies such as the World Bank and UNESCO more critically and constructively. Moreover, through this study, I have developed a deepened and more nuanced understanding of the relationship between education and development, as well as the use of LSAs in developing countries. Finally, this research has opened up my professional career as an academic researcher and has provided inspiration for future research on the related themes.

Acknowledgements

This research experience has been a journey for me. I owe a great debt of gratitude to the many people who have helped me to arrive where I am today.

In particular, this thesis could not have been commenced without the inspiration of Dr. Euan Auld. To you, dear friend, my sincere thanks. I am most grateful to my principal supervisor, Professor Paul Morris, for his guidance, support and immense knowledge. I have spent the four most wonderful years of my life under Professor Morris' supervision and have learned much more than simply how to research and write a thesis. Thanks to his monthly seminars, I have also had the unique opportunity to connect with an amazing group of intelligent people who have been a great source of inspiration and academic and spiritual support. I give special thanks to Dr. Ai Lian Chee, Dr. Fei Yan, Dr. Maryam Mohamed, Dr. Annette Bamberger, Dr. Sue Grey, Dr. Laura Oxley, Dr. Heather Kinuthia, Heela Goren and Yan Kin Cheung.

I express my heartfelt gratitude to Professor Robert Cowen for his continuous support, wisdom and generosity. I have enjoyed every conversation I have had with him since embarking on this journey. His advice for treating writing as a piece of art and playing the game of editing have greatly influenced how I perceive academic research. I am privileged to have had the opportunity to work with him. I also express my sincere gratitude to Professor Robert Adamson for his insights and constructive feedback and for all the time and knowledge he has contributed to my research.

I would like to give special thanks to Dr. Helen Longlands for her kindness and encouragement over the years. Her persistence opened up an opportunity for me to complete an important interview. Special thanks goes also to Dr. Jeremy Rappleye, who shared his knowledge on the work of OECD; to Professor Elaine Unterhalter for putting me in touch with a contact in one of the PISA-D participating countries; to Professor Christian Ydesen for sharing his work and resources on the history of OECD; to Dr. Christine Han for her generous support with my ethical applications; and, to Professor Terri Kim for providing comments on my presentations. I would also like to acknowledge the invaluable contributions of all my interview participants.

I am greatly indebted to my family in China for their unconditional love, care and support. I thank the China Scholarship Council for providing the financial support which made this study possible. Last but not least, I express gratitude to the Chinese Embassy

in London for providing opportunities to help facilitate two major China-UK ministerial-level meetings, where I gained insights into the China-England Maths Teacher Exchange programme launched in 2014 as a result of Shanghai's outstanding performance on PISA 2009 and 2012.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Changing the global education agenda

The year 2015 was significant for the relationship between education and international development. It saw the end of an approach to development based on the vision of the Millennium Development Goals (MDGs) which – with Goal 2 and 3 – prioritised educational access and was associated with the maxim of Education for All (EFA). That vision was replaced by the Sustainable Development Goals (SDGs). In particular, Goal 4 shifted the focus from access to quality and identified minimum standards of quality for low- and middle-income countries to be delivered by 2030. This new vision is oriented towards ensuring 'learning for all'.

The rationales for this post-2015 shift of the global education agenda are twofold. The first is that expanding access alone would not be sufficient for realising education's promise as the main driver of development. The United Nations Educational, Scientific and Cultural Organisation (UNESCO) – the UN's lead agency for the Education 2030 agenda – highlighted that as an important lesson to learn from the MDGs and declared that education systems need to ensure all children and youth are learning and acquiring relevant skills:

> The focus on education quality, learning and skills highlights the important lesson: the danger of concentrating on access to education without paying enough attention to whether students are learning and acquiring relevant skills once they are in school. … For countries and communities that embrace the need to bring quality education to all, the benefits are enormous. (UNESCO 2015, 6)

The second is that there is an urgent need for action in addressing the quality problem, given that millions of children are still not learning and large learning gaps prevail:

> Despite significant progress since 2000, an estimated 59 million children of primary school age and 65 million adolescents of lower secondary school age – of whom girls remain the majority – were still out of school in 2013. In addition, many of those in school are not acquiring basic knowledge and skills. At least 250 million primary-school-aged children, more than 50% of whom have spent at least four years in school, cannot read, write or count well enough to meet minimum learning standards. (ibid., 36)

The situation was described as a global 'learning crisis' by prominent international agencies working on education and development, such as the World Bank, UNESCO and UNICEF (the United Nations International Children's Emergency Fund). Recently, to highlight this crisis, the World Bank (2019a) introduced the concept of 'learning poverty',

referring to the inability of children to read and understand a simple text by the time the children were ten years old. This concept draws on new data developed in joint effort with the UNESCO Institute for Statistics (UIS) (World Bank 2019b) which estimates that 53% of all children in low- and middle-income countries suffer from 'learning poverty'.

In parallel, a further lesson drawn from the experience with EFA and the MDGs was that 'business as usual' will not bring quality education to all. To improve learning, it has to be measured and tracked. While those prominent agencies adopt different approaches with regard to what quality is and how it can be measured (see Unterhalter 2018; Wulff 2019), they generally reach the consensus that progress on learning has to be based on the evidence of data and given enhanced visibility (see also Gorur 2019).

For example, UNESCO (2015, 25) stated that 'building on the lessons of EFA and the MDGs, states should invest in and scale up innovative, evidence-based and cost-effective approaches that enable all individuals to gain access to, participate in, learn through and complete a quality education'. This call was echoed in a UIS blog which likens education ministers to 'air traffic controllers' who simply do not have the data to help them navigate through the 'storm' of the global learning crisis (Crouch and Silvia 2018). Similarly, the World Bank's *World Development Report 2018: Learning to Realise Education's Promise* suggested concrete policy steps to help developing countries address this crisis by stronger learning assessments, using evidence of what works to guide education decision-making and to mobilise societal and political support for education reform.

Although learning assessments, especially international large-scale ones, can be costly for countries with scarce resources and limited data infrastructure, the counter argument given is that the cost of *not* assessing learning outcomes can be much higher. In its report, UIS (2016b, 15) calculated the cost of international assessments and compared it with government expenditure in public education per student, warning that 'it is very expensive to **not** assess education, with the consequent risk of not providing students with the skills demanded by the labour market and [sic] increasing inefficiencies' (original emphasis). The report thus urged the need 'to leverage resources in the most efficient manner possible and rely on the guidance provided by information on learning outcomes' (ibid.). To this end, UIS has been working on identifying common minimum proficiencies as a measure of learning, based on a harmonisation of existing international, regional and national assessments. However, not all are satisfied with the database

developed at UIS. UNICEF (2018, 59), for instance, claimed that 'the data are not internationally comparable and do not provide sufficient data points to establish trends'.

This particular contemporary debate then is fairly clear. There is a contemporary crisis – a perceived 'learning crisis' – and there is increased reliance on learning metrics to resolve it. However, it may be suggested that the concept of 'quality' is defined narrowly, and it can also be suggested that education is seen as a means to achieve economic ends, and that all that seems to be needed is more careful collection of data and evidence. As Unterhalter (2019) hinted, this overlaps with the institutional histories of these actors in developing learning metrics. This has become a 'common sense' orthodoxy, or a 'mainstream narrative', as Dahlberg, Moss and Pence (2013, vii) call it. Thus, one of the starting points of this thesis is: what are the politics, sociology and history of this stance and its implications in terms of the OECD?

Secondly, the theme will be pursued that a general consensus about 'quality' and how to see and measure it provides an enabling environment for the Organisation for Economic Co-operation and Development (OECD), to push forward its policy agendas armed with strong statistical and technical capacity in assessing learning outcomes. The OECD also has strong persuasive power, for example, in utilising the media to help disseminate assessment results (Grey and Morris 2018). The traditional UN agencies that have a mandate in global education, including UNESCO and UNICEF, are in a less advantaged position given their relatively lower level of expertise and experience in developing international large-scale assessments (ILSAs). The World Bank is an exception, as it has successfully evolved into a knowledge bank (Mehta 2001; Mundy and Verger 2016; Klees et al. 2020); it has also formed a strategic partnership with the OECD (Auld, Morris and Rappleye 2019).

Among others, Tikly (2017) depicts the EFA movement and the formulation of SDG 4 as a form of governance, in which key actors are involved in a process of building and maintaining legitimacy. This perspective is helpful for understanding the monitoring work of SDG 4: for some traditional actors like UNESCO, it can be read as central to the struggle over the control of SDG 4 monitoring; whereas for the OECD, it is more about carving out a new market, consolidating its established role, and adapting its approaches to become optimally suited to the context:

> The adoption of the 2030 Approach and the Sustainable Development Goals (SDGs) by the international community is an important driver for OECD's work on sustainable development. As a hub of expertise, experience, and innovative ideas, the OECD provides tools, frameworks, data

and knowledge, working constructively and inclusively with partners and stakeholders to support the 2030 Agenda to the best of its ability. (OECD 2015, n.p.)

Thus part of this thesis will be to focus on the strategic work that the OECD has undertaken to adapt to the post-2015 era while also shaping and extending the niche it has evolved into the field of education and international development. Primarily, the organisation has promoted its most successful policy instrument – the Programme for International Student Assessment (PISA) – for usage in low- and middle-income countries as its major contribution to SDG 4; and the programme used to promote PISA is called PISA for Development (hereafter PISA-D).

OECD contribution to SDG 4: from PISA to PISA-D to a global learning metric

PISA, a standardised assessment of the knowledge and skills of 15-year-olds, is a framework for international comparisons. By assessing how well young people perform against their peers in readiness for life and work, PISA compares the quality and effectiveness of education systems they are situated within (OECD 1999). Since its inception in 2000, PISA has been primarily focused on skills in the key subjects of reading, mathematics and science. The major domain was reading in 2000 and 2009, mathematics in 2003 and 2012, and science in 2006 and 2015. In addition to the major domain, in 2012 the OECD included one-off components, such as a computer-based assessment of problem solving and of financial literacy. In 2018 it added an assessment of global competency.

While expanding what can be measured, PISA has also enlarged its coverage of territories: from 28 OECD member countries and 4 partner countries in PISA 2000 to 34 OECD countries and an additional 30 partner countries (and economies) in PISA 2012. By and large, these are relatively affluent territories; the initial design of PISA instruments was aimed to serve the needs and priorities of the OECD countries (OECD, 1999; see also Chapter 4). Thus, alongside the expansion of PISA in terms of scale and scope (Sellar and Lingard 2014) are discussions and debates about its relevance and usage in low- and middle-income countries. These can be observed, for example, from India which withdrew from PISA in 2009 after performing poorly and claimed that the test had not been sufficiently adapted to the Indian context (Edwards 2019). Kyrgyzstan was the only low-income country which engaged in PISA in 2006 and 2009. The results were

extremely poor for Kyrgyzstan – taking the last place, and it stopped participating in the 2012 cycle (Shamatov and Sainazarov 2015).

Scholarly research has highlighted the significant challenges that low- and middle-income countries face in participating and benefiting from PISA (e.g. Bloem 2013; Lockheed, Prokic-Breuer and Shadrova 2015). The main challenges identified were the lack of financial resources and institutional capacity, and less relevant results due to a non-representative sample of 15-year-olds and the clustering of students at low proficiency levels. These challenges are far more significant across low-income countries, where the majority of 15-year-olds are either not in school or have repeated a grade (see Chapter 6). This means that these groups of youth are not qualified to sit the PISA test as it is school- and age-based. Consequently, the results failed to provide meaningful information on equity as the most disadvantaged are most likely to drop out or repeat a grade.

Recognising these challenges, the OECD launched PISA-D in 2013. It claimed that PISA-D would enhance PISA survey instruments and data-collection methods to produce results which better reflect the realities in low- and middle-income countries but which are still comparable with those of the main PISA; it also claimed that through participation in PISA-D, countries would develop their capacity in managing ILSAs and using the results to inform national education policies and improve learning (OECD 2013c, 2013d). To achieve these, PISA-D focused on three technical strands:

> Strand A – Composing test items to target at the lower levels of performance as measured on the PISA scale (i.e. Level 2 or below).
>
> Strand B – Improving contextual questionnaires to capture the diverse situations in low- and middle-income countries.
>
> Strand C – Establishing methods to incorporate the out-of-school youth into the assessments.

The OECD decided that the majority of PISA-D test items would remain at level 2 or below, because on the PISA scale, Level 2 is considered as the basic proficiency level 'at which students begin to demonstrate the competencies that will enable them to participate effectively and productively in life as continuing students, workers and citizens' (OECD 2018c, 9). It also wanted to connect PISA-D with SDG 4. Target 4.1 aims at ensuring that all children complete free and compulsory education leading to at least a minimum proficiency level of literacy and numeracy. The OECD was aware that this target would need to be measured and tracked on a global scale, and so positioned PISA-D as an important contribution (see Chapter 5). Thus, while the OECD has been working on

enhancing PISA to capture a wider range of contexts and provide a granular analysis of students' performance at the lower end of the PISA scale, it also claimed that the results from PISA 2021[1] cycle onwards (i.e. 2024, 2027 and 2030) would provide a consistent measure of progress on this target.

Therefore, a theme explored in this thesis is the role that PISA-D played to expand the OECD's global influence. The OECD envisioned that, through PISA-D, the number of countries participating in PISA will reach 170 by 2030 and thus OECD will become a truly global organisation (Ward 2016). It has acknowledged PISA-D in its Strategy on Development endorsed in May 2012 as a critical step to take forward the Organisation's global expansion. The following statement is illustrative of the goals which PISA-D aims to achieve:

> PISA for Development is well placed to support global efforts to frame a learning goal in the context of the post-2015 agenda and to provide a single universal metric for measuring progress towards this. PISA is regarded as one of the most important education policy instruments in the world today and there is a compelling logic for making this relevant to developing countries so that greater numbers of these can benefit from the surveys and analysis. The project therefore addresses squarely the core elements of the Strategy on Development, particularly the adaptation of the most successful OECD policy instruments to make these more relevant for developing country contexts. (OECD 2014a, 49)

In other words, this thesis will suggest that the aims attached to PISA-D are multiple and they are interlinked. PISA-D is an enhanced framework for global comparisons (i.e. to enable comparison and identify 'best practices' on a global scale); a development strategy supporting the SDGs (i.e. to provide a major source of data for monitoring SDG 4 and to help improve education quality in developing countries); and, a core element of the OECD's Strategy on Development (i.e. to extend policy instruments to developing country contexts and to become a global organisation).

These features indicate that PISA-D differs from the OECD's other testing instruments, such as comparative assessments of higher education learning outcomes (AHELO), of adult skills (PIAAC), of teaching and learning (TALIS), of early years (International Early Learning and Child Well-Being Study), and of schools (PISA-based test for schools). These assessments are mainly practised in affluent countries with relatively higher levels of performance and the technical capacity to use the results. In

---

[1] Due to the global COVID-2019 pandemic, the OECD has recently announced to postpone the PISA 2021 assessment to 2022 and the PISA 2024 assessment to 2025 to reflect post-COVID difficulties. PISA 2022 will focus on mathematics, with an additional test of creative thinking.

addition, these assessments also involve minimal levels of collaboration and competition between the OECD and traditional UN agencies that have a mandate in global education. While UNESCO, along with UNICEF, is officially leading and coordinating efforts on framing and monitoring SDG 4, the dissemination of PISA-D sends out a clear message that PISA will serve as the universal monitoring tool for SDG 4 and in so doing position the OECD as a key player in the field.

This introductory narrative begins to frame the thesis. These introductory themes are formalised and made explicit in the next sub-section.

Research aims and questions

Although there is a small body of literature recognising the potential impact of PISA-D on the OECD's role in global educational governance post-2015 (Addey 2017; Kaess 2018; Auld, Morris and Rappleye 2019; Gorur, Sørensen and Maddox 2019; Addey and Gorur 2020; Li and Auld 2020), there is a lack of studies systematically interrogating its origins, practices, and outcomes.

This thesis addresses that need. It aims to achieve this through three main lines of inquiry. These are elaborated below.

First, the significance of PISA-D within the OECD Strategy on Development suggests the need for an historical reading of PISA-D. This means that PISA-D is seen as not a singular event but a process that PISA is undergoing. The process is about both continuity and change, essential for understanding the OECD's contemporary and emerging governance approach in comparative education and international development. An historical perspective is thus necessary to situate PISA-D within an analysis of the OECD's changing agendas and approaches to education and development from their inception.

While there is an extensive body of literature focusing on the OECD's educational activities, these have been primarily concerned with a specific aspect, such as the success of PISA and its impact on national policies. Little work has focused on the OECD's much longer history and sociology of involvement in education, which has markedly changed over the years and continues to evolve. Leimgruber and Schmelzer (2017, 5) concurred, noting that highlighting a specific aspect of the organisation's work may shed light on its distinctive mode of governance, but it 'impedes a more thorough understanding of the

OECD's role', especially among postwar multilateral organisations. The result – they suggest – is a rather static view of the OECD. Similarly, Ydesen (2019, 2) states that 'while most research recognizes the enormous importance of the OECD as a global education policy shaper, little effort has been made in gaining a better understanding of the developments and events that made it possible for the OECD to assume this dominant role'.

There are a few notable exceptions, including the work of Eide (1990) and Papadopoulos' (1994) which have provided detailed historical accounts of the OECD's educational activities from the 1960s to the 1990s. In particular, Papadopoulos (1994) offers insights into the ways in which the OECD's educational activities had endeavoured to both respond and contribute to broader policy trends and developments that had to do with significant social, economic and political changes.

However, this thesis will extend this analysis to incorporate more recent developments, and suggest that the features of PISA-D are distinctive from the previous forms and the geopolitical and economic context within which the OECD is situated has also changed significantly. The extended analysis will then allow me to trace the origins of PISA-D and identify how these historical circumstances, events, and/or processes have shaped its incarnation into the contemporary form. In this way, by linking the past to recent developments, and internal dynamics to external realities, the analysis can contribute to a more thorough understanding of the OECD's changing role. It can also provide a solid foundation for examining the related practices and outcomes. The first line of inquiry is thus guided by the following question:

What are the origins of PISA-D and how have they shaped its contemporary form?

Having traced the origins of PISA-D within an understanding of its historical, institutional and ideological basis, the second line of inquiry explores the OECD's practices in promoting and legitimating PISA-D globally. We still have little knowledge of how PISA-D has been strategically woven into the SDGs as a legitimate source for monitoring SDG 4. Addey (2017, 11) suggests that 'this enigma is one of the many unknown stories that occur in policymaking and global education agenda setting – a dimension that would greatly benefit from further empirical research'. At the level of global institutions, although the OECD was previously involved in the MDGs, that focus was primarily on aid and development co-operation under the work of the Development Assistance Committee (DAC) (OECD 1996a, 2011e). With regard to the post-2015

education agenda centred around SDG 4, the OECD is a relatively new player, compared to more traditional UN agencies such as UNESCO and UNICEF (Coleman and Jones 2005; Mundy 2007; Mundy and Manion 2015; Elfert 2018).

Furthermore, while existing research recognises the OECD's intention to expand into low- and middle-income countries through PISA-D, the focus of these studies is placed on global-local entanglements. For example, Gorur, Sørensen and Maddox (2019) examine how global PISA standards get translated into locally relevant PISA-D practices in Zambia, and Addey and Gorur (2020) recently added analysis of translation practices in Ecuador and Paraguay. Addey and Sellar (2018) identify the multiple purposes which drive countries' participation in ILSAs such as PISA-D. In a recent paper, they further argued that these diverse purposes should be incorporated into evaluating the benefits of ILSA participation on a broader scale (Addey and Sellar 2019). By contrast, Auld, Morris and Rappleye (2019) analyse the arrival of PISA-D in Cambodia, highlighting the 'imperialist' tendency of the OECD's approach towards these countries (see also Kaess 2018). By and large, little can be found which explains how PISA-D is promoted and legitimised and how the OECD has manoeuvred itself into a key position in the context of the SDGs.

Lingard (2017) suggests that the OECD has been strategic in using media coverage to reinforce the political impact of PISA on education ministers and public perceptions. However, as Grey and Morris (2018) demonstrate, the role of media in that process is much more complex, because the media operates on its own logic and thereby exerts a powerful influence on the OECD in a different way. Grey and Morris instead identify this kind of influence as a form of 'mediatised governance'. In the case of PISA-D, the strategic use of media to explain how the OECD has promoted and generated support for PISA-D is also problematic in that the long-awaited PISA-D results were released 'quietly' in December 2018. This was in marked contrast to the usual media coverage that accompanies the release of the OECD's other assessment results, such as PIAAC (see Hamilton 2018) and the recent PISA global competency results. This was also apparent from the *Report on the Implementation of the OECD Strategy on Development* (2014a) where there was no mention of media usage among the 'main deliverables' of this project.

Morgan (2017) attributes the legitimacy of PISA to the OECD's 'marketing and policy framing practices' which help to promote and distribute PISA and create demand among countries and policy makers. These practices remain relevant. However, the

legitimation of PISA-D would require extra efforts from the OECD due to the past, mainly negative, PISA experience in some low- and middle-income countries (e.g. India and Kyrgyzstan) and the significant challenges that countries face in engaging with PISA (Bloem 2013; Lockheed, Prokic-Breuer and Shadrova 2015). Moreover, to promote PISA as the universal monitoring tool for SDG 4, the OECD would also need to engage with the traditional UN agencies that are officially leading and coordinating the monitoring process.

Based on these considerations, it is necessary to analyse the diverse range of PISA-D activities and documents, which will help to understand how the OECD accomplished these tasks. In this way, the analysis will shed light on the 'enigma' proposed by Addey (2017) and the processes by which the OECD managed to establish a clear role for itself in the context of the SDGs. The analysis thus addresses the following research question:

How has PISA-D been promoted and legitimated, and what strategies have been employed by the OECD to establish its role in monitoring SDG 4?

The third line of inquiry examines the PISA-D outcomes, including the adapted assessment instruments, the findings from school-based assessments, and recommended policy interventions. These outcomes will be examined with reference to the challenges and goals which the OECD claimed PISA-D would respectively overcome and deliver.

This aspect of inquiry connects to wider research into the rise of ILSA phenomenon in low- and middle-income countries. However, the thesis does not seek to add to the already substantial literature on the 'why' question: why do countries participate in ILSAs? The answers can be easily found in the works of, for example, Kamens and McNeely (2010), Kamens (2013), Kijima (2013), Pizmony-Levy (2013), Valverde (2014), Lockheed (2015), Lockheed, Prokic-Breuer and Shadrova (2015), and Addey and Sellar (2018, 2019). While Kamens (2013) and Ramirez, Schofer and Meyer (2018) insist on the power of the global testing 'culture' in transcending national boundaries of different types of countries, scholars such as Valverde (2014) and Steiner-Khamsi (2017) tend to highlight the 'agency' of often weak, poor, or underachieving educational systems.

The focus of this inquiry is on the 'how' question: how do ILSAs help to address the learning needs or challenges faced by low- and middle-income countries? Although ILSAs have received enormous attention in recent decades, Lockheed and Wagemaker

(2013, 297) observe that 'the processes whereby they have been developed and used [in developing countries] have received significantly less attention, even though attention to such processes may be salient for improving education policy'. Understanding these processes is also significant in that the issues in low- and middle-income countries are unique and warrant systematic research (Wagner, Wolf and Boruch 2018). This means, in addition to the lack of financial resources and institutional capacity in these countries, there are also concerns about the learning needs of the marginalised populations (Wagner and Castillo 2014), about the efficacy of ILSA instruments to be applied in under-resourced settings (Maddox 2014, 2015; Scherman, Bosker and Howie 2017), about the validity of the assessment results, as well as about their implications for comparing countries with diverse social-economic, cultural and linguistic conditions (Cowen 2014; Crossley 2014; Kaess 2018; Auld, Li and Morris 2020).

The thesis aims to examine PISA-D outcomes in the light of these concerns. Specifically, it asks how PISA instruments have been adapted to address the challenges that low- and middle-income countries face in engaging in ILSAs, how the assessment results are generated and compared, and what policy insights they provide for the PISA-D participating countries. The answers to these questions are vitally important given that the assessment results will be used to track learning progress on SDG 4 Target 4.1.1(c) (UIS 2019). Moreover, they will inform education policies as part of the strategy to support countries' future development (i.e. PISA *for* development). As the OECD (2018c, 13) claimed, the results are meant to provide 'policy makers [in these countries] with data and evidence that can be used to determine what they can do to improve their educational systems and, ultimately, ensure that their students obtain the skills needed to succeed in tomorrow's world as set out in the Education SDG framework'. Last, by answering these questions, the thesis aims to enrich the literature which focuses on the application and use of ILSAs in low-resourced settings. The third line of inquiry is pursued through the following question:

> *How has PISA-D addressed the challenges that low- and middle-income countries face in engaging in ILSAs, how are the assessment results generated and compared, and what policy insights do they provide for the piloting countries?*

Research contributions

Some potential contributions of this study are indirectly illuminated in the preceding section. Overall, there has been limited research on PISA-D due to its recent occurrence (officially launched in 2013). The thesis, therefore, engages with the existing literature which details the arrival of PISA-D in piloting countries (Auld, Morris and Rappleye 2019) and which delves into the pilot process itself (Gorur, Sørensen and Maddox 2019; Addey and Gorur 2020). It also recognises that the drivers of countries' participation in PISA-D are multiple and sometimes what is claimed may not respond to what is actually used (Addey 2017, 2019; Addey and Sellar 2018, 2019). The thesis also aims to build on and extend the literature which raises concerns and criticisms about the extension of PISA into low- and middle-income countries, including, for example, the role of PISA-D as a source of global educational governance (Addey 2017; Auld, Morris and Rappleye 2019; Li and Auld 2020) and as a practice of educational colonialism and imperialism (d'Agnese 2015; Tikly 2016; Brock-Utne 2018; Kaess 2018; Silova and Auld 2019).

However, what has been missing from existing research is a more thorough understanding of PISA-D, especially of where it comes from, how it is promoted and legitimated, and what outcomes are generated from the 'pilot' process. Thus the initial contribution of this thesis is to fill the research gap by developing an in-depth analysis of PISA-D based on the three main lines of inquiry identified.

An in-depth analysis also advances the potential for critical analysis of the OECD's role in global educational governance post-2015. This is the second contribution of the thesis. It aims to enrich the literature which seeks to understand the OECD's changing role in global educational governance (Henry et al. 2001; Rizvi and Lingard 2006; Martens and Jacobi 2010; Sellar and Lingard 2014; Auld and Morris 2018; Ydesen 2019; Lewis 2020). In the post-2015 context, this has culminated in the work of PISA-D and related post-pilot arrangements (Addey 2017; Auld, Morris and Rappleye 2019; Li and Auld 2020). The purpose of this thesis, however, is not to validate a single theoretical perspective. It aims to deepen the understanding of the OECD's governance by analysing the nature and modus operandi of its approach to education and development of which PISA-D is a critical element. To achieve this goal, the thesis draws on a range of perspectives and insights from other fields and disciplines, which will be described in the following chapter.

Third, the thesis also aims to contribute to research which examines the role and impact of ILSAs on improving the quality of learning. Generally, there are two major

approaches to the measurement of learning. One is the prominent global approach which focuses on national averages and international (or regional) comparability. This approach is represented by large-scale assessments (LSAs) and is preferred by the key international agencies. They believe that LSAs are critical both for monitoring and for achieving the SDGs (e.g. OECD 2016c; UIS 2016a, 2016b; UNICEF 2018; World Bank 2018). The other approach is more localised and oriented towards variations within countries and across constituencies. This latter approach has greater potential for addressing the learning needs of individuals from particular groups, such as those who are left behind and who are ethnic-linguistic minorities (Wagner 2011; Wagner and Castillo 2014; Vaughan 2015; Wagner, Wolf and Boruch 2018).

While PISA-D responds to the global approach that allows for universal assumptions and international comparability, it also claims to support the learning needs of out-of-school youth, specifically through Strand C to develop instruments to approach and assess this group of population. By examining the outcomes of PISA-D on this strand, the thesis aims to facilitate understanding of how the global approach can help to improve children's learning and development. In other words, the thesis contributes to understanding to what extent (or not) PISA is for 'development'.

Thesis structure and outline

The thesis has seven chapters including this introductory chapter. Chapter 2 reviews the literature and Chapter 3 describes the methodology. Chapters 4 to 6 are the analysis chapters and are followed by Chapter 7 as the discussion and concluding chapter. A detailed introduction is provided at the beginning of each subsequent chapter. Below I outline the general content of these chapters.

Chapter 2 will situate the study within three major but interrelated bodies of literature: comparative education (CE), international relations (IR), and education and international development (EID). The chapter will first draw on critical research traditions from CE to understand PISA-D as a framework for enabling international comparisons. This feature responds to the 'applied' form of CE that is well rehearsed in the literature. The CE literature locates the philosophical and methodological grounds for this particular form and provides further insights into policy interventions promoted through PISA-D. The chapter will then draw on IR scholarship to understand the changing role of IOs in global

governance. It will review the three major approaches to the study of IOs: realism, neoliberalism, and constructivism. It finds that the sociologically informed constructivist approach is particularly relevant to interpret the OECD's changing role and the extension of PISA into low- and middle-income countries. The constructivist approach sheds light on the feature of PISA-D as a core element of the OECD's Strategy on Development. The chapter will also review international development studies to understand the use of PISA for development. It will focus on two major aspects of relations: PISA for economic growth and PISA for improving learning.

Chapter 3 will describe the research process and explain my adoption of the critical hermeneutic approach. It will revisit the paradigmatic traditions (including hermeneutics, phenomenology, and critical theory) which this critical hermeneutic approach contains and rests upon, and how they are helpful for designing my own research to answer the major research questions. The application of the critical hermeneutic approach will be supported with three data sources: organisational documents, interviews, and literature. I will read these sources intertextually to elicit meaning and develop 'situated' knowledge of the OECD and PISA-D. The interpretive process will follow a hermeneutic 'circle–spiral' (Schwartz-Shea and Yanow 2012), which is a continuous interpretive path moving towards a deeper and richer understanding. This implies that the concluding chapter will be more reflexive and not intended for closing this study.

Chapter 4 will seek to understand the origins of PISA-D and the processes by which they have contributed to the contemporary form. I will proceed by analysing the OECD's evolution from the 1950s to the post-2015 era (i.e. 2015-2030). The interpretation will mainly draw on Barnett and Finnemore's (2004) work on international organisations to illuminate the OECD's highly evolved capacity and the expansionist nature of its approach to education and development. PISA-D in this regard is critical for the OECD's expansion and engagement with the SDGs (in particular SDG 4). The chapter will also allude to the OECD's emerging governance approach – a more comprehensive approach to education and development. In education, this manifests in composing a 'larger' PISA programme and proselytising a 'data revolution' in low- and middle-income countries. The chapter will interpret that the ambition for the OECD's educational governance post-2015 is to achieve assessment for all.

Chapter 5 will identify and analyse the strategies deployed by the OECD to promote PISA-D as a legitimate learning metric. It will first focus on the ways in which the OECD

relates to and establishes partnerships with UNESCO and UNICEF to gain socio-political acceptability in defining and measuring SDG 4. It will then unpack how the OECD strategically reframes the global education agenda to weave in PISA-D, presenting PISA as the most relevant, competent, and demand-driven learning metric. Following these, the chapter will elicit the meaning of PISA-D in relation to a 'global PISA community', which the OECD constructs to provoke the benefits of participation and a strong sense of membership and belonging. The chapter will finally interpret the series of PISA-D technical meetings, workshops, and seminar as a specific form of OECD marketing strategy – which I term as PISA 'boosterism'. Moreover, the chapter will reveal how the PISA-D countries are construed as backward – and as needing to learn from other countries participating in PISA.

Chapter 6 will examine the major outcomes of PISA-D. These will include the modified assessment instruments based on the three technical strands, the findings from school-based assessments, and the effective policies promoted in the main OECD report and specific country reports. The chapter will also reflect on the OECD's strategic use of 'pilot' to help legitimate but not evaluate the project. My focus on the instruments will be both to understand and to highlight problems associated with (i) the linkage between PISA-D and PISA test items to enable international comparisons and benchmark basic learning, (ii) the adoption of the Education Prosperity Approach (EPA) to help enhance the contextual questionnaires, and (iii) the design and delivery of the out-of-school component to include children not in school. The second part of this chapter will illustrate how the OECD compares student performances and interpret them through a range of contextual factors prescribed by the EPA. It will show that the OECD's interpretation stabilises the explanatory logic already established within the EPA and is not derived from the actual data. The last part of the chapter will focus on the policy actions proposed in the reports, illustrating how these are problematic, and how they serve to ensure countries' continual involvement in PISA.

Chapter 7, as the concluding chapter, will begin with revisiting the previous chapters, focusing on how these are connected and intertwined. Afterwards, I will summarise my contributions to knowledge in relation to the following themes: global educational governance, organisational legitimacy, and banal imperialism. Considering the global COVID-19 pandemic, I will provide additional observations on the OECD's recent COVID-related educational initiatives. My observations will show how these initiatives

continue to exclude out-of-school youth; this neglect, along with the OECD's failure and unwillingness to tackle the real limitation of PISA for low- and middle-income countries, undermines its newly gained moral authority as well as its expertise. Finally, I will reflect on the research limitations in terms of scope and theorising, and then propose areas for future research.

Chapter 2. Literature Review: PISA as 'Applied' and PISA for 'Development'

Introduction

In Chapter 1, I identified the distinctive features of PISA-D based on the main purposes it serves to achieve. The first feature is its role as 'an enhanced framework enabling international comparisons'; the second feature is its position as 'a core element of the OECD's strategy for development'; and the third is its portrayal as 'a development strategy supporting the SDGs'. Understanding these features provides insights into the OECD's contemporary and emerging governance approach to education and development, and is, therefore, central to this thesis. In this chapter, I review the extensive literature on these features. The literature review will have two functions: the first is to position myself within relevant bodies of scholarship; and the second is to draw critical insights for interpreting the origins, practices, and outcomes of PISA-D.

The thread running through this review is the understanding that PISA-D is problematic for low- and middle-income countries for three reasons. First, much of the comparative education (CE) literature suggests that PISA-D, as a framework for comparing all types of countries with vastly different conditions, embodies an 'applied' form of CE. However, this 'applied' form of CE is built on flimsy philosophical and methodological grounds. Second, constructivist international relations (IR) scholars argue that international organisations (IOs) as bureaucracies are key players in constructing and spreading this particular form of comparison. PISA-D in this respect is critical for the OECD's expansionist agenda. Third, international development studies show mixed results for the positive impact of PISA on economic growth and on children's learning. This indicates that the OECD proclamation of PISA-D as a development strategy supporting SDG 4 is also problematic.

Consequently, the structure of this review will consist of three major parts. These are summarised and presented in Figure 2.1. The first part (shown on the left of the figure) is a review of the CE literature which will help to understand the first distinctive feature of PISA-D. In particular, by contrasting it with the academic traditions of CE, I locate the underpinning ideas and assumptions associated with this specific form. I perceive that there are at least three levels of challenge (i.e. philosophical, methodological, and postcolonial) which PISA-D confronts as it enrols low- and middle-income countries on

to the existing global comparison framework. The review will then revisit the two major theoretical approaches (world culture theory vs. local agency/meanings) which help to explain the rise of the ILSA phenomenon in general and the global extension of PISA more specifically.

The second part of the review (shown at the bottom of Figure 2.1) will incorporate IR theories and the research on global governance. I will first draw on IR theories to read the role of IOs and the ways in which they operate. There are three major approaches: realist, neoliberalist, and constructivist. I find that while realism and neoliberalism are helpful for understanding the OECD's role and functions in the earlier decades, they are inadequate for explaining its dramatic change in the subsequent periods. Instead, constructivism based on Weberian arguments appears fruitful for making sense of the OECD's changing role in global governance. To better understand the specific ways in which the OECD exercises power in education, I will then review the global governance research, focusing on the OECD's operating mechanisms (i.e. idea production, policy evaluation, and data generation). These will be read along with the different forms of authority, i.e. rational-legal, delegated, moral, and expert, which constructivist scholars identify as the major sources of power of IOs. I also propose that recent research on organisational legitimacy provides a useful framework for interpreting the dynamics of the authority of IOs and the OECD strategies, in particular, to establish its role in monitoring SDG 4. Moreover, in this part of the review, I will also touch on the emerging global education policy (GEP) studies which assess the nature of global drivers affecting educational change. While both IR and GEP address the role of IOs and the ways in which they operate, the GEP studies provide further insights into the capacity of educational service providers to influence the work of IOs.

The third part of the review (shown on the right of Figure 2.1) will explore the education and international development (EID) literature. This EID literature will help to understand the extension of ILSAs (PISA in particular) into low- and middle-income countries as a development strategy supporting SDG 4. Specifically, the review will focus on two dimensions: national economic growth and student learning. The first dimension will be assessed through studies that examine the economic impact of student achievement on PISA informed by a human capital reframing of education (located within development economics). I will consider this dimension by reflecting on the limitations to the cognitive-economic model which, however, underpins the global extension of PISA

and its claims for development. The second dimension will be based on developmental psychology to understand the educational impact of ILSAs. This will involve examining the fundamental aspects of large-scale survey designs (including educational constructs, sampling methods, and contextual variables) which affect the validity, reliability, and practical feasibility of the assessment approach. The second dimension will highlight the acute challenges that PISA-D needs to address in order to be relevant to the contexts of low- and middle-income countries. Moreover, the review will extend beyond an appraisal of technical issues to encompass critical reflection on how PISA-D 'contributes to' the global knowledge divide and serves to advance the OECD's global expansion.

Figure 2.1 A schematic overview of the literature review: understanding PISA and PISA-D for low- and middle-income countries.

An 'applied' form of CE:

- **A scientific approach to comparison**
  - assumptions: *commensuration* and *comparability*.
  - features: *generality* and *causality, prescriptive* and *policy-oriented*.
  - practices: *identify* 'what works' and *transfer* best policies.
- **PISA-D problematic for three considerations**
  - philosophical, methodological, postcolonial.
- **Two major theoretical approaches**
  - *global testing culture* transcends national boundaries.
  - focusing on *local meanings* to interpret the global.

Extension of ILSAs into low- and middle-income countries

PISA & PISA-D

A development strategy supporting SDG 4:

- **Impact of PISA on economic growth**
  - human capital reframing: *achievement* vs attainment.
  - *cognitive-economic* model: improved PISA scores lead to future economic growth.
  - problems: *modernisation* theory – narrow vision, linear progression, statistical 'othering'.
- **Impact of PISA on student learning**
  - large-scale survey design: validity and reliability – constructs, sampling, contextual variables.
- **Two additional related perspectives**
  - knowledge *colonialism*: intercultural, North/South divide.
  - *new imperialism*: univocal; PISA all-encompassing.

A core element of OECD governance post-2015

- **IR theories on the role of IOs**
  - *realism*: instruments of powerful states.
  - *neoliberalism*: structures for international cooperation.
  - *constructivism*: purposive actors with immediate interests; four types of authority: rational-legal, delegated, moral, expertise.
- **OECD governance mechanisms in education**
  - *mechanisms:* idea production, policy evaluation, data generation.
  - *conditioning factors*: external players, internal dynamics, policy constraints.
  - *an* authority in SDG 4: rational-legal, expert.
  - not *in* authority: delegated, moral – seek to *gain* to monitor SDG 4.
- **Organisational legitimacy in the GEP field**
  - dynamics: interplay of different types of authority.
  - legitimacy management: gaining, maintaining, and repairing.

PISA-D as an 'applied' form of comparative education

A scientific approach to comparison

PISA, as a form of ILSA, is at its core an exercise in comparative education (CE). To be more specific, it is an 'applied' form of CE premised on the idea that, based on careful measurement of educational processes and outcomes, the resultant comparisons can generate hard data and evidence that can be used to identify 'what works' and to transfer them (from one place) to improve education systems (in another place). Morris (2015) identified the implicit assumptions that underpin this idea. These include: first, the educational processes and outcomes are commensurable; second, the measurement is valid and consistent; third, the causes of high performance can be isolated from the original context; and fourth, policies based on those causes can be transplanted into a different context. Retrospectively, this idea of developing an 'applied' CE was recognised when Marc-Antoine Jullien first attempted to define and systematise the field in 1817. Since then, 'many of the comparative educationalists of the 1960s such as C. Arnold Anderson, Edmund King, Brian Holmes and Noah and Eckstein signed up to different versions of an ideology of science' (Cowen 2014, 291).

An alternative perspective to this is the academic field of CE that is more culturally, historically, and philosophically rooted. This perspective is promoted in the works of Issac Kandel (1955), Nicholas Hans (1958), Robert Ulich (1964) and Michael Sadler (1990). For this group of scholars, CE is mainly explanatory and interpretive. It is 'not a predictive or a policy-oriented or practical/applied social science' (Kazamias 2009, 39). The following statement from Sadler (1990, 50) provides a classic illustration of how we gain insights into a foreign education system

> …If we propose to study foreign systems of education, we must keep our eyes on the brick and mortar institutions, nor on the teachers and pupils only, but we must also go outside into the streets and into the homes of the people, and try to find out what is the intangible, impalpable, spiritual force which, in the case of any successful system of Education, is in reality upholding the school system and accounting for its practical efficiency…; In studying foreign systems of Education we should not forget that the things outside the schools matter even more than the things inside the schools, and govern and interpret the things inside.

Sadler's statement points to understanding 'the intangible, impalpable, spiritual force' which upholds the education system. The primary focus of this intellectual inquiry is thus less on generality and causality but oriented towards understanding and interpreting the invisible forces that are of essential value. The academic field of CE, as Mattheou (2009,

63) puts it, does not seek to 'agree on the pragmatic dimensions of the field, on its potential to provide sound advice to policy makers'.

Although the separation of CE between these two major perspectives remains an enduring theme, the rise of the ILSA phenomenon, as represented by PISA, has added popularity to the applied form of CE. This tendency can be observed from two aspects. One is that politicians are using PISA data (as a source of evidence) to help legitimate national education policy preferences (Grek 2009; Morris 2012; Ozga 2012; Waldow 2012; Auld and Morris 2014). The other is that researchers are adopting comparative approaches to help obtain additional resources and symbolic advantages (Nóvoa and Yariv-Mashal 2003; Cowen 2006, 2007; Roberts 2007; Crossley 2014). Both aspects are achieved by way of comparison. Consequently, the academic field of CE is being altered, becoming increasingly 'absorbed' and 'reshaped' by the political conditions for the production of academic knowledge (Cowen 2014).

PISA-D problematic: philosophical, methodological, postcolonial

While this 'applied' CE has, to date, primarily involved more affluent countries, it is now expanding into low- and middle-income countries via PISA-D. The expansion is celebrated by those promoting 'applied' CE (such as the OECD and the World Bank) as it allows for a wider range of comparisons and, therefore, greater diversity of sources for promoting international policy transfer. However, from the academic CE point of view, such an expansion is problematic based on at least three levels of consideration.

The first is philosophical in that the expansion raises further doubt about the basis for applying education indicators to generate comparisons. For example, Nóvoa and Yariv-Mashal (2003) highlight the importance for understanding comparison as a 'historical journey' in which the focus of comparison is not on education 'facts' and 'realities' but on problems – problems that carry a history and gesture towards different possible futures. They argue that the task of educational comparatists is to multiply space and unfold time to embrace new zones of looking rather than subscribe to the 'applied' thinking.

The second related problem is methodological as the expansion means a higher level of difficulty in comparing countries with vastly different conditions. Indeed, Wagner and Castillo (2014) question the relevance and feasibility of applying a specific set of indicators to measure learning across different societies. They suggest that

> [if] the primary goal is comparability, less attention may be paid to the local and cultural validity of the definitions and classifications of learning. Further, the data may be less meaningful and potentially less applicable at the local level. This is a natural and essential tension between universalistic *etic* and context-sensitive *emic* approaches to measurement. (Wagner and Castillo 2014, 63)

This tension becomes more acute in the case of PISA-D because the conceptual framework of PISA (i.e. DeSeCo) and the questionnaires were aimed at serving the objectives and priorities of the OECD countries (see Chapter 4).

While the OECD claims that PISA-D would enhance the relevance of PISA instruments to the situations in low- and middle-income countries, it remains unclear how this could be achieved. The 'translation' process (Addey and Gorur 2020) involves the contradictory task of reconciling global standards and local contexts in order to establish international comparability (see also Rottenburg et al. 2015). Gorur, Sørensen and Maddox (2019) observe that the 'translation' from PISA into PISA-D in Zambia has encountered issues with standardised booklet design, challenges of sampling across cultures, and difficulty in locating and reaching out-of-school youth. Although adjustments and compromises are both made to achieve a thin layer of relevance, the authors (2019) suggest that local contexts are 're-contextualised' to comply with global PISA standards.

The methodological challenge also relates to the third aspect of consideration that is postcolonial (discussed further in Part 3). Aside from preserving distinct cultures and traditions, many developing countries are endeavouring to liberate themselves from the shackles of their colonial masters (Unterhalter 2009). In Africa, for example, Brock-Utne (2016, 40) notes that educational researchers 'have been constantly debating what quality in education may mean in their own context and how it should be assessed'. She suggests that, in order for Africa to build up an education system that adheres to their own values (e.g. care for others and cooperation), 'it is critical both to teach in the languages learners speak and understand and to avoid the Western testing regime' (2016, 41). For many scholars like Brock-Utne, the purpose of education and assessment is closely linked to identity and self-reliance rather than glide into the society of the 'international' spectacle (Nóvoa and Yariv-Mashal 2003).

Two major theoretical approaches: the global and the local

There are two major theoretical perspectives in the comparative literature that help to explain the rise of the ILSA phenomenon. First, the expansion of PISA seems to conform

to the trajectory of a 'global testing culture' which transcends national boundaries and local realities. Scholars in the world culture or world society traditions, such as Ramirez, Schofer and Meyer (2018) posit that the rationale for ILSAs is to gauge and promote learning, driven either by the dominant economistic discourses or the rights-based discourses. At the fundamental level, Ramirez, Schofer and Meyer (2018, 357) suggest that it is the 'technocratic visions of social improvement' and the authority of science that enable ILSAs 'a universalistic transnational reach'. Second, in contrast to this 'global testing culture' is the scholarly focus on 'local meanings' of ILSA phenomenon (e.g. Schriewer and Martinez 2004; Steiner-Khamsi 2004; Schriewer 2012). These scholars insist on the versatility of ILSAs in serving diverge ends, highlighting the central agency of policy makers in appropriating ILSAs for national agenda setting. For example, Steiner-Khamsi (2017, 444) suggests that the strength of PISA lies in its elusiveness and polyvalence – meaning different things to different policy actors depending on the context – so much so that 'focusing on the local to understand why the global resonates'.

Although these two prominent perspectives provide insights into the expansion of PISA into low- and middle-income countries, Carney, Rappleye and Silova (2012, 367) contend that they have reached a theoretical impasse,

> …both sides of the debate offer complementary perspectives in explaining the interaction between the global and the local: world culture theorists are charged with documenting global phenomena while their opponents content themselves with analyzing their local variations. What was once a lively debate has thus fallen into a comfortable set of complementarities about educational convergence.

Moreover, within this debate, the role of IOs is also conceived differently on each side. For world culture theorists, the expansion of PISA is a relatively homogeneous institutional form spreading around the world, and so the OECD's role is mainly perceived as a 'carrier' of pre-structured norms rather than a purposive actor with its own interests. Whereas for scholars focusing on local meanings and appropriation, engagement in PISA is mainly demand driven – an asset which state actors could trade upon for political and/or economic gains. In this line of thinking, while the powerful influence of the OECD is acknowledged, it depends very much on the mediating role of the state.

Therefore, both sides share no interest in the changing role of IOs as autonomous sources of power, which foster a global educational culture and, at the same time, shape local political decisions and preferences. From this, the thesis resonates with Resnik's

call (2006, 173) that it is time to bring IOs back in, 'which means treating international organisations as agents of the global educational arena and as producers of a global educational culture'. To achieve this, the second part of the review (summarised and presented at the bottom of Figure 2.1) will draw on literature from IR and research on global governance to better understand how IOs are perceived and the ways in which they operate. The central focus is on understanding the OECD's changing role in global educational governance.

Bringing IOs back in: OECD's changing role in global educational governance

Realism, neoliberalism and constructivism

Three major approaches from the field of international relations (IR) have contributed to the study of IOs. These are realism, neoliberalism, and constructivism.

In conventional realist thinking, the world order is fundamentally shaped by the relations between nation states, each focusing on pursuing its own interests and power (Carr 1964; Holsti 1985; McKinlay and Little 1986). As Carr (1964, 10) puts it, realists tend to 'emphasize the irresistible strength of existing forces and the inevitable character of existing tendencies, and to insist that the highest wisdom lies in accepting, and adapting oneself to, these forces and these tendencies'. Given that nation states have been the major international actors for so long, IOs do not have independent standing in international politics. Instead, they are more likely to be used by the more powerful states to take advantage of the less powerful, and so the role of IOs is seen as being at the service of powerful states. Accordingly, IOs change and adapt their agendas and approaches in response to the needs of their most powerful member states (Morgenthau 1967; Zacher 1979; Gilpin 1981; Viotti and Kauppi 1987).

While realism was criticised later for being state-centric, 'it seemed to provide a useful framework for understanding the collapse of the post-World War I international order in the face of serial aggressions in the Far East and Europe, World War II, and the Cold War' (Holsti 2004, 5). In the case of the OECD's work on education during the Cold War period, the realist perspective is helpful for understanding the great power dynamics behind its major activities. For instance, technical and scientific competition between the Soviet Union and the U.S. generated strong strategic interest in information sharing and

standard setting in education, which led to the first education work undertaken by the OECD and some of the earliest bilateral funding for cross-national testing and comparison (Papadopoulos 1994; see also Mundy 2007). However, considering that powerful states dominate IO activities, the realism approach seems to contradict the observation that educational ideas and experiences within the OECD came to be strongly influenced by the policy ideas and practice of the small Nordic countries (see Eide 1990). Moreover, it is also inadequate for interpreting the ensuing changes in the OECD's approach and its growing influence in global educational governance. These suggest that in addition to state powers, there are other potent forces at play in international politics.

Neoliberalism takes realism's state-centric approach a step further by focusing on the mechanisms of international cooperation. It argues that nation states cooperate in common strategic or instrumental interests, particularly when they hold common values or face uncertainty (Milner 1992; Abbott and Snidal 1998). In this regard, neoliberal scholars emphasise the importance of institutions and regimes as potential causative factors in international politics (Krasner 1982; Keohane 1984, 1988). Based on Krasner's (1982, 185) definition, regimes are 'sets of implicit or explicit principles, norms, rules and decision-making procedures around which actors' expectations converge in a given area of international relations'. By acknowledging these important factors, neoliberal scholars came to recognise the possible independent effects of IOs on nation states. For example, Abbott and Snidal (1998, 9) suggest that IOs can 'shape understandings, influence the terms of state interactions, elaborate norms, and mediate or resolve member states' disputes'. By and large, however, the neoliberalism approach insists that states are the principal actors and that they create IOs primarily to facilitate cooperation and reduce transaction costs in an increasingly interdependent world (Rittberger and Zangl 2006).

Although this neoliberal perspective captures the OECD's facilitator role in the period before the 1990s (as explained in Chapter 4), it is still not sufficient for interpreting how the role changed dramatically in the following decades. In other words, neoliberal scholarship has not been able to recognise IOs as autonomous sources of power, and therefore, it is inadequate for explaining how IOs perform after being created and how they evolve and expand with the passage of time. Another shortcoming is that this perspective tends to share a rosy view of IOs as doing 'good', as purveyors of progress and enlightenment. In the case of the OECD, this seems to run into empirical anomalies as, for example, the OECD has also triggered the politicisation of sectors, such as

economics, planning and education, within member states (see the 1st period of OECD activities; Chapter 4). As far as PISA-D is concerned, this perspective seems to reflect the position that the OECD holds for extending PISA into low- and middle-income countries: PISA for development. By drawing on development studies literature, I will show later how such a positioning is problematic.

In contrast to realism and neoliberalism, constructivism goes beyond the state-centric approach to IOs by focusing on the normative or ideational factors that define state interests and shape international politics (Finnemore 1996a, 1996b; Finnemore and Sikkink 1998; Ruggie 1998; Barnett and Finnemore 1999). In particular, constructivist scholars such as Murphy (1994) and Barnett and Finnemore (1999, 2004) adopt a perspective from sociology, conceiving IOs as social creatures or bureaucracies. More specifically, they draw on Weberian arguments about the normative power of the rational-legal authority which bureaucracies embody and with which they produce and control social knowledge. In addition to rational-legal authority, Barnett and Finnemore (2004) identify that IOs also derive authority from other sources, notably from their moral understanding, expertise, and delegated tasks. Therefore, the perception of IOs as bureaucracies provides the basis for constructivist scholars to treat them as ontologically independent actors rather than as merely structures for state interaction.

By emphasising the political agency of IOs, constructivists also offer a different conception of the relationship between IOs and their environment. On the one hand, IOs teach states about the value and utility of their missions, thereby shaping state policies and structures (Finnemore 1993, 1996a). For instance, Resnik (2006) shows how the OECD has taught states and even other organisations to think about education and the economy in new ways. Similarly, Rubenson (2006, 2015) demonstrates how the OECD framed lifelong learning in relation to economic and social production functions, which has led to changes in states' approach to adult learning and education reforms. On the other hand, Barnett and Finnemore (1999, 704) note that "[o]rganizations respond not only to other actors pursuing material interests in the environment but also to normative and cultural forces that shape how organizations see the world and conceptualize their own missions'. This applies, for example, to UNESCO. Elfert (2018, 204) observes that UNESCO struggles to hold on to its humanistic position in an environment which favours a more technical and economistic approach to education and had to take a 'technical turn' in order to build up its statistical capacities in monitoring the EFA initiative.

In the case of the OECD, the latter can be observed from its endeavour to adapt educational activities to respond not only to the economic imperative but also to emerging social and environmental concerns (Henry et al. 2001; Gorur 2014; Sellar and Lingard 2014; Rubenson 2015; Auld and Morris 2019; Rappleye et al. 2019; Li and Auld 2020). A strategic part of PISA-D is to include out-of-school youth in the assessment to provide some information on equity. This helps the OECD to conceptualise its mission of extending PISA as an effort to support SDG 4 that leaves no one behind. As I illustrate below, this kind of changing dynamic relates to the different types of authority that IOs possess or seek to possess in the field of global governance. Understanding the interplay between them provides insights into the OECD's changing role in the global education arena and how it also relates to other prominent IOs.

From IR theories to the global governance research

According to Mundy (2007, 342), the 'constructivist turn' in IR scholarship has predicated a newly restructured debate around global governance.

Traditionally, the term 'global governance' was associated with the need for regulating interdependent global relations in the absence of an overarching political authority (Robertson 2014). However, the collapse of the Soviet Union has fundamentally altered the terms of debate. As Overbeek et al. (2010, 698) observe, 'with a global alternative out of the way, it became possible to promote a depoliticized and watered-down version of "global governance" as the ideal consensual and non-adversarial manner to manage the world's affairs'. Primarily because globalisation and technological developments have transformed many aspects of social life, state capacity and sovereignty in managing domestic affairs have also been diminished. Processes of economic integration and neoliberalisation not only required a curtailed role for governments but also created non-state actors, such as transnational corporations (TNCs) and non-governmental organisations (NGOs), that help deliver state services. These factors led to a reframing of the world order as a system of 'governance without government' (Rosenau and Czempiel 1992).

According to Rosenau (1992), the world system is sustained and regulated through a myriad of activities which unfold at three basic levels: (i) the ideational level of what people normally perceive; (ii) the behavioural level of what people routinely do; and (iii) the institutional level where governance occurs. Specifically, the ideational level of

activity 'involves the mental sets, belief systems, shared values, and any other attitudinal or perceptual screens through which the events of world politics pass before evoking reactions or inactions' (1992, 14). The behavioural level of activity consists of what actors do 'in a regular and patterned way' according to their ideational understandings. For example, the behavioural patterns could be about conformity, negotiations, or resistance to 'recurrent behaviors that are so salient as to shape and reinforce the prevailing conceptions of the underlying global order' (ibid.). The institutional level of activity then involves the more formal and organised dimension of the prevailing order, such as institutions and regimes. More empirically driven constructivist IR scholars engage in the governance debate by examining the roles played by global norms, IOs and other non-state authorities in those activities (Mundy 2007).

Indeed, the OECD had played a critical role in the global shift from government to governance. On the one hand, it promoted what Overbeek et al. (2010) call 'a depoliticized and watered-down version of global governance' in education. This manifested in its work of developing the DeSeCo (*Definition and Selection of Competencies: Theoretical and Conceptual Foundations*) programme between 1998 and 2002. DeSeCo guided the development of PISA and other OECD assessments. It has served as the overarching 'ideal' framework for managing and monitoring education systems within OECD countries ever since the late 1990s (see the 2nd period of OECD activities; Chapter 4). On the other hand, the OECD, along with other supranational organisations (such as the World Bank and the International Monetary Fund [IMF]), helped disseminate neoliberal education reforms across the globe (known as the global education reform movement [GERM]). These reforms have not only led to the privatisation and commodification of public education services, but also engendered the terrors of performativity which transform individual education practitioners into self-enterprising subjects with constant struggles (Ball 2003).

Both of these aspects suggest that the OECD has evolved, with the passage of time, into an autonomous, influential source of power. In the context of globalisation and the gradual 'thickening' of global governance (Robertson 2014), its influence reaches not merely the supranational level but also inward towards nation states and subnational groups. Realism and neoliberalism appear limited in explaining such dramatic change, let alone capture the dynamics of organisational legitimacy in the global education arena (or GEP). Below, I draw on the constructivist approach to global governance research to

better understand the ways in which the OECD operates and exerts its influence. This small body of literature is also critical for understanding the meaning of PISA-D as a key element of the OECD's strategy for development and, therefore, the OECD's educational governance post-2015.

OECD governance mechanisms and legitimacy management

The OECD's changing role in global educational governance has prompted research on its governing approaches (e.g. Martens 2007; Grek 2009; Woodward 2009; Sellar and Lingard 2014; Addey 2017; Williamson 2016; Auld, Morris and Rappleye 2019; Ydesen 2019).

At a general level, Martens and Jakobi (2010) distinguish three mechanisms which the OECD applies to convert its primarily non-binding activities into effective policy influence. These are idea production, policy evaluation, and data generation. They suggest that the OECD is a regular producer of discourses. The legitimacy of OECD discourses rests on the expertise of its staff who are able to identity issues of current and future relevance and claim causative influences. By raising and defining issues in specific ways, the OECD directs policy discussions and provides possible solutions. In terms of policy evaluation, the OECD mainly uses the mechanism of peer review to generate pressure and public scrutiny. Martens and Jacobi argue that through this mechanism, 'the OECD not only defines the relevant issues but can also draw up and disseminate recommendations on actions in its member states in order to address previously identified problems' (2010, 9). This process depends heavily on the OECD's technical capacity to generate large quantitative and comparative data sets (Martens 2007). From the constructivist IR perspective, creating the appearance of 'depoliticization' is essential for IOs to wield authority with regard to their recommendations, and 'quantification vastly enhances the power of these claims of objectivity and impartiality' (Barnett and Finnemore 2004, 24).

However, Martens and Jakobi (2010) suggest that the effects of the OECD governance mechanisms are conditioned by three major factors, namely external players, internal dynamics, and policy constraints. They explain that

> (1) … External players are, for example, other international organizations or specific countries that initiate policy development and respective governance mechanisms within the OECD. (2) Internal dynamics, such as the structure of the organization, the influence of specific member states, or the competition of different departments, can also influence the choice of governance mechanisms. (3) Policy constraints relate to the fact that different policy fields display different

> dynamics depending on how contested or consensual they are, and thus have an impact on how the OECD exercises governance. (Martens and Jakobi 2010, 7)

Indeed, these factors are considered in the OECD's *50th Anniversary Vision Statement* (2011c), which underscores the need for the organisation to evolve into a more effective and inclusive global policy network. To this end, it commits to a new paradigm for development, based on a more comprehensive approach, to deepen partnerships with developing countries and other IOs externally and to achieve greater synergy across different directorates internally. The focus of this approach is to account for the heterogeneity of growth and development models and the differentiated institutional settings and capacity to tap into resources and address any binding constraints (OECD 2012d). These aspects provide an important background for understanding the strategic role of PISA-D within the OECD's broader agenda in global governance post-2015, which I elaborate in Chapter 4.

In the context of the SDGs, the OECD co-produces the 'globalized education policy discourses' (Lingard and Rawolle 2011). This is evident, for example, in the construction of a global 'learning crisis' as noted earlier. As Chapter 5 interprets, the discourse is primarily used to prepare for identifying minimum learning standards and promoting PISA as the legitimate learning metric to measure those standards. As such, the OECD discourses are more than merely self-perpetuating (Sellar and Lingard 2014); they are also productive in the sense that the OECD gains moral authority from them – by urging those working in the development industry to use PISA to gauge progress on SDG 4. The latter aspect responds to Eccleston's (2011, 248) point that an IO's political authority 'is at its zenith when its rational-technical agenda aligns with prevailing social values and sentiments'.

However, in terms of SDG 4 monitoring, it is worth reiterating that the OECD is not delegated to perform those tasks. The UIS is the mandated agency to lead the definition and monitoring of SDG 4 targets. As Barnett and Finnemore (2004) note, while expertise contributes to the authoritative status of an IO, delegation is central to conferring authority on the IO. The IO in authority is the IO occupying the role or position society recognises as legitimate to exercise power. This implies that the OECD can be seen mainly as being *an* authority rather than being *in* authority regarding the work on SDG 4.

Research on organisational legitimacy, however, has come to recognise legitimacy as an entity that is constantly being negotiated, built and maintained (see Buchanan and

Keohane 2006; Bitektine 2011). This is because legitimacy results from 'a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed systems of norms, values, beliefs, and definitions' (Suchman 1995, 574). Taking this sociological constructivist position, Suchman suggests that 'organisations *can* and *do* formulate strategies for fostering legitimating perceptions of desirability, propriety, and appropriateness' (ibid., original emphasis). The process relies on the communication between the organisation and its various audiences, which 'extends well beyond traditional discourse, to include a wide range of meaning-laden actions and non-verbal displays' (ibid., 586). In other words, the skilful management of legitimacy requires a diverse arsenal of techniques.

From this instrumental/managerial standpoint, and drawing on Suchman's definition, in Chapter 5, I interpret the legitimation strategies which the OECD employs to gain authority in monitoring SDG 4. The interpretation can be read along with Edwards et al.'s (2018) recent work which focuses on UNESCO repairing its legitimacy through EFA coordination and the production of the Global Monitoring Report (GMR). Both analyses contribute to understanding the dynamics of organisational legitimacy within an emerging GEP field which refers to the international political space in which powerful agencies interact and compete to promote their educational discourses and preferred policy solutions (Verger, Novelli and Altinyelken 2012, 2018; Mundy et al. 2016). Within these processes, private institutions and policy entrepreneurs are becoming more influential based on different forms of ideas and technical expertise (Ball 2012; Junemann and Ball 2015; Menashy 2016). Compared to previous decades, these private actors now 'have more spaces and opportunities to influence IOs without the necessary mediation of the state' (Verger, Novelli and Altinyelken 2018, 16). This thesis recognises the 'competence' of PISA-D contractors for fulfilling the contradictory task of reconciling global PISA standards and local specificities. In Chapter 6, I will examine the specific work achieved (and not achieved) to ensure the implementation 'success' of PISA-D.

IR scholarship and the global governance literature, therefore, provide a rich body of works to understand the processes by which the OECD has come to wield global authority. Ontologically different from the realist attitude of treating IOs as instruments of powerful states and neoliberalism as arenas for state interaction, the constructivist approach to global governance research 'looks in new ways at how forms of international authority are socially constructed and historically contingent, rather than materially or

historically fixed' (Mundy 2007, 351). Given that norms and ideas are critical factors in the structuring of world affairs, IOs are key producers and teachers of them (Finnemore 1993; Barnett and Finnemore 1999, 2004). Moreover, IOs are also legitimacy managers, as they compete to extend or, as in the case of UNESCO, restore their influence and power in the field of global educational governance.

Although the OECD was initially created to primarily serve the interests of the U.S., over time it has evolved into an autonomous, influential source of power. This ontological independence suggests that the OECD is a purposive actor and PISA-D a purposive activity. The purpose of PISA-D is closely aligned with the OECD's strategy for development, which envisages a globally inclusive OECD based on a more comprehensive approach to education and development. PISA-D is thus 'a core element of the OECD strategy' to maintain its relevance and to continue the expansion.

To obtain legitimacy in monitoring SDG 4, the OECD has to be a skilful 'manager' in order to foster legitimating perceptions of PISA-D's 'desirability, propriety, and appropriateness' (Suchman 1995). Key to its success is the sense of moral type of authority of PISA for 'development'. In what follows, I draw on the third body of literature from international development studies to understand how PISA is associated with development and to what extent it is a contributory factor. These empirical studies provide further insights into how to interpret the OECD discourses presented in Chapters 5 and 6.

PISA-D as a development strategy supporting SDG 4

Empirical studies on the impact of PISA on economic growth

Discussions on the use of ILSAs for development often revolve around the reframing of education as an investment in human capital to promote economic development. The conventional human capital theory (HCT) asserts that the education level of a country is an important factor for explaining national economic growth (Schultz 1961, 1971; Becker 1964). Because, unlike traditional factors (such as labour and land), human capital – pertaining to the knowledge and skills that workers have acquired as a result of education – contributes to productivity and thereby to earnings at an individual and collective level. The increase in human capital stock thus leads to economic growth.

Education is considered as an investment from this perspective, and its impact on earnings and other related outcomes (such as health) is measured using the 'rate of return'. For example, Mincer (1974) identifies that an additional year of schooling increases lifetime earnings by between 8 to 15 percent. Drawing on Mincer's regression model, Psacharopolous (1994) and Psacharopolous and Patrina (2004) summarise empirical evidence for the patterns of the returns on schooling at different levels across developing countries. They further suggest that the highest returns are recorded for low- and middle-income countries.

While early estimates have mainly used school attainment (or years of schooling) as measures of human capital, more recent studies have turned the focus to cognitive skills and to the use of ILSA scores to arrive at more accurate measures of human capital. In particular, economists Eric Hanushek and Ludger Woessmann have introduced PISA test-score measures into growth regressions, arguing that the achievement measure is substantially more positively associated with economic growth than the attainment measure. For instance, according to Hanushek and Woessmann (2008, 638),

> [a]fter controlling for the initial level of GDP per capita and for years of schooling, the test-score measure features a statistically significant effect on the growth of real GDP per capita in 1960–2000. According to this specification, test scores that are larger by one standard deviation (measured at the student level across all OECD countries in PISA) are associated with an average annual growth rate in GDP per capita that is two percentage points higher over the whole forty-year period.

For a long time, international development agencies have pursued the expansion of schooling as a primary component of development (Mundy 1998, 2006). The evidence that Hanushek and Woessmann provide seems to speak to the significance of cognitive skills and of improvement and measurement of learning outcomes. This partly explains why the key agencies noted earlier, in addition to extending or maintaining their influence and power, have come to prioritise the use of ILSAs to gauge learning and drive development (e.g. UIS 2016a, 2016b; The Global Partnership for Education 2018; World Bank 2018).

Indeed, the cognitive-economic imperative underpins the OECD PISA. Relating back to the CE literature, Auld and Morris (2014) note that contemporary versions of applied CE (as represented by PISA) can be distinguished from their earlier forms primarily by their focus on measuring student performance to explain educational and economic outcomes. The result has been that 'earlier uncertainties about how education works, how it impacts society, and how to best allocate scarce resources are being quickly replaced

by contemporary certainties that raising test scores will result in higher levels of economic growth (GDP)' (Komatsu and Rappleye 2017, 1). Thus, for future economic growth in low- and middle-income countries, the OECD projects the rates of return as follows:

> If Ghana could educate all of its students to at least the basic level of reading and mathematics skills [as measured by PISA], it would see a gain over the lifetime of children born today that, in present value terms, is 38 times its current GDP … For lower-middle income countries, the gains would be 13 times current GDP and would average out to a 28% higher GDP over the next 80 years. And for upper-middle-income countries, whose students generally perform better academically, it would average out to 16% higher GDP. (Schleicher 2018, 139)

These projections, however, have been widely criticised. A detailed review of the analyses is in Komatsu and Rappleye (2017). Overall, they categorise the existing critiques into two strands. The first strand focuses on the temporal mismatch within the data (Ramirez et al. 2006; Kamens 2015) and the second strand questions the extrapolation of the relationship identified in historical data to make future projections (Klees 2016; Stromquist 2016). Extending these critiques, Komatsu and Rappleye (2017) re-examined Hanushek and Woessmann's original data sets in order to challenge the validity of their claims (that improvements in ILSAs such as PISA will lead to higher levels of economic growth). Their results showed that (i) the relationship between test scores and economic growth was not consistently strong over time, and (ii) the relationships between the change in PISA scores and corresponding change in economic performance in the subsequent period 'were unclear at best, doubtful at worst' (ibid., 5). Recently, the authors added PIAAC studies to further refute the causal relationship established between cognitive levels and GDP growth rates (see Rappleye and Komatsu 2019). Another recent study by Feniger and Atia (2019) shows the opposite effect that student performance on ILSAs appears to respond to economic changes rather than predict them. Similarly, a new working paper by Patel and Sandefur (2020) offers strong evidence that country revenue and family income are equally important factors for predicting test scores.

While this cognitive-economic model based on international educational performance data has been much debated and critiqued in the academic literature, it continues to gain momentum. In the development field, Klees (2016, 644) comments that it 'has been ubiquitous and widely accepted as an important mechanism for educational planning, evaluation, and policy making'. The lack of ILSA infrastructure in many low- and middle-income countries has provided the compelling logic for the OECD to extend PISA into these countries, for 'development'. As the name itself suggests, PISA-D is a

deliberate effort of the OECD to position PISA as a development strategy supporting SDG 4. Similarly, Addey (2017) observes that the OECD attempts to extend and consolidate its influence by aligning the aims of PISA and PISA-D with its strategy for development and SDG 4. These efforts signal the OECD's grander vision for shaping the future of education and development, which Auld, Morris and Rappleye (2019, 5) identify as including the following key components:

> (i) a universal standard of education quality (PISA); (ii) a cognitive-economic model of education, with quality defined in terms of human or knowledge capital (PISA as the proxy); (iii) assessment as a public good, or 'human right' (merging human rights and human capital); (iv) alignment (or replacement) of national-level and regional assessments with a global standard (PISA); (v) a global policy network, mentorships, and transfer of 'best practices'; (vi) increased private involvement at each stage of the process; (vii) incentivized compliance, and punitive accountability (i.e. outcome-based lending initiatives such as Disbursement Linked Indicators).

These studies provide critical insights into understanding the meaning of PISA-D in relation to the OECD's global governance post-2015. Building on the existing analyses, I explain in Chapter 4 this emerging, more comprehensive governance approach – and how it operates through mechanisms of configuration and modulation. Before moving to a review of the effects of PISA on student learning, I summarise below the major limitations of this cognitive-economic model that are important for interpreting the implications of the extension of PISA into low- and middle-income countries as a development project. These limitations are considered by drawing on perspectives from modernisation theory and postcolonial scholarship.

Cognitive-economic model: narrow vision, linear progression, statistical 'othering'

First, the cognitive-economic model narrows our perception of development significantly. According to this approach, the factors that matter for development are the cognitive skills acquired in formal school settings. It assumes economic growth as the only or primary feature of development, and consequently, downplays other influences (such as health and resources), other dimensions of learning (e.g. awareness, curiosity), and its variations across distinct cultural and linguistic settings. All of these are important aspects of human development. To put it another way, by relating education to economic development, the cognitive-economic model focuses more on the 'capital' rather than on the 'human' aspect of development. It is 'not daring' to make the choice of understanding the child as 'a rich child' endowed with infinite capabilities and different possible futures, as Dahlberg, Moss and Pence (2013, 7) would put it. Another more comprehensive view

of the relationship between education and development is provided by McCowan and Unterhalter (2015, 7):

> Here, education becomes constitutive of development, with the line between the two increasingly hard to draw. We can see this relationship in the participatory approaches to development emerging from the 1970s involving community empowerment, in Freire's radical ideas on pedagogy, and in the emphasis on agency in the human development and human rights paradigms, in which human beings' emancipation and participation in collective decision-making – and consequently learning and personal development – are central.

Second, given that the cognitive-economic model defines progress primarily through the prism of test scores, it fosters the legitimating perception of PISA as the basis for improvement. Consequently, the message for countries that have not engaged in PISA is to *participate* so that they are able to embark on the development journey, and for those that are currently involved but at the bottom of the PISA rankings, their priority is to *import* 'best practices' and to reform education systems accordingly within domestic contexts. The trajectory is essentially placed on the PISA developmental scale. The process is described by the OECD (2006, 18):

> The beginning [for countries which successfully addressed the challenges of quality, equity and efficiency] lies in accepting international benchmarking in educational performance as a basis for improvement ... It is only through such benchmarking that countries can understand relative strengths and weakness of their education systems and identify best practices and ways forward. The world is indifferent to tradition and past reputations, unforgiving of frailty and ignorant of custom or practice. Success will go to those individuals and countries which are swift to adapt, slow to complain and open to change. The task of governments will be to ensure that countries rise to this challenge.

The OECD-prescribed development journey is rooted in modernisation theory which involves transforming the 'traditional' into the 'modern'. A well-known approach to this is the U.S. economist Walter Rostow's (1960) five stages of development which is a linear progression model of economic development that envisages countries advancing through the stages to arrive at the modern self-sustaining state. In education, the implication for low- and middle-income countries is that just as in Western Europe and North America, mass schooling should underpin the formation of the modern state also in the newly independent countries of Asia and Africa (McCowan 2015). The underlying logic, encapsulated in Rostow's (1960, 19) statement, is that 'man need not regard his physical environment as virtually a factor given nature and providence, but as an ordered world which, if rationally understood, can be manipulated in ways which yield productive change'. Similar to Rostow's concept, the contemporary cognitive-economic model follows the rationalist approach to social engineering and planning (see also Elfert 2019),

assuming that features of 'best practices' are seen to apply to all countries, and that progress for those that are lower down on the PISA developmental scale involves them becoming more like those that are higher up (see Chapter 6).

Third, as a consequence of the standpoint described above, PISA-D can be read as *continuing* the process of 'othering' (Fischer 2017) by statistically placing low-performing developing countries remote in space and behind in time (Silova and Auld 2019; Auld, Li and Morris 2020). While the world order based on the Western model has been well maintained, Silova and Auld (2019, 276) observe, from the perspective of international educational performance data that, certain East Asian countries (such as China, Singapore, Japan, and South Korea) continuing to top global educational rankings clearly upsets the 'unspoken order of things'. They note that the West no longer appears to be the single powerful model for the rest of the world to emulate, and so 'the OECD's world order – and its careful positioning of countries along with the East/West or South/North axis – was suddenly destabilized' (ibid.). It is at this very moment when well-established, taken-for-granted foundations become breached and when the 'script writers' step back in to restore the status quo that we explicitly observe the normally implicit order of things.

Drawing on this insight, I interpret in this thesis that a part of the 'repair' work is taken up by PISA-D so that (i) it maintains the PISA technical standards to ensure international comparability, (ii) it adds 'easier' test items to enrol low- and middle-income countries on to the development journey, and (iii) it applies/prioritises the Western developmental model (i.e. the educational prosperity approach [EPA]) to adjust the contextual questionnaires and interpret the findings.

However, as I will also argue, the sense of belonging and identity is not explicitly linked to being part of the Global West (or Global North) and (or) a Western civilisation as Silova and Auld (2019) suggest. It is articulated in relation to being part of the *global* PISA community in which all types of countries come to sit together and emulate each other. The community operates based on the mechanisms of peer-learning and mutual accountability, and with deviation defined in relation to the OECD *average* and (or) *top-performing* or *fast-improving* countries in the PISA rankings, such as the East Asian countries, as well as Brazil, Peru and Vietnam (see Chapters 5 and 6). As touched upon earlier, PISA-D represents the OECD's emerging governance approach which accompanies its grander vision (see Chapter 4). This emerging approach is much more

inclusive and flexible, and potentially more digital and individual, compared to the previous forms of mechanisms discussed in the global governance literature.

Empirical studies on the impact of PISA on student learning

ILSAs have greatly impacted on how we think about and act upon education, but little is known about how exactly they have influenced teaching and learning practices and contributed to our understanding of learning. Kanjee (2018, 79) concurs, noting that the primary value of ILSAs is to provide information for use in policy change but 'whether or how this "benefit" translates to practice depends on a range of factors and varies substantially between and within countries'. With regard to low- and middle-income countries, the factors of financial resources and institutional and technical capacity, as well as the political and cultural environment, can affect the effective use of ILSA information (Lockheed 2012; Bloem 2013; Lockheed, Prokic-Breuer and Shadrova 2015). Moreover, in some low-income countries, engagement in ILSAs could be driven by donors' interest in obtaining data for accountability needs, or as in the case of Cambodia (Auld, Morris and Rappleye 2019), agencies' desire for influence, rather than by local demand. There are certainly multiple reasons that drive countries' engagement in ILSAs (Addey and Sellar 2018; 2019) and, therefore, their actual use of data.

However, even for countries with substantial resources, such as the U.S., ILSAs 'based on average national scores or current achievement gaps encourage policy-makers to draw misguided conclusions about school improvement and may actually lead them to pursue inappropriate and even harmful reforms' (Carnoy and Rothstein 2015, 124). Carnoy and Rothstein's point relates to ILSA designs which focus on national averages and international rankings. The advantage of large-scale survey designs relies on the scale of data collection which allows for greater generality and/or comparisons over location and/or time (Vaughan 2015). However, the disadvantage in generating average scores and league tables is that they only reflect certain aspects, such as examination achievement, as opposed to assessing the actual quality and equality of education. Because of the scale of data collection and the claims of generality, it is important to scrutinise the fundamental aspects of ILSA designs, including the skills to be assessed (i.e. the development of educational constructs), the population sampling methods, and the contextual and demographic variables (Braun and Kanjee 2006; Wagner 2010, 2011). Wagner (2018, 233) perceives that each of these aspects 'is tied to a set of assumptions

and compromises, and each will influence the validity, reliability and practical feasibility of the chosen assessment approach'.

ILSA designs: constructs, sampling methods and contextual variables

Educational constructs refer to the ultimate targets of measurement (e.g. literacy and numeracy). They are usually derived from theoretical considerations and are organised into a framework which is later translated into a set of test specifications (Braun and Kanjee 2006). A better understanding of educational constructs and their theoretical underpinnings requires resorting to the literature exploring the connection between assessment and learning theories. As Baird et al. (2017, 317) suggest, 'if assessments are to serve the goals of education, then theories of learning and assessment should be developing more closely with each other'. From the development perspective, these areas of research are also relevant in that they speak to the 'psychology of development' (Wagner 2018, 22). Psychologists have long asserted that development is largely a cognitive and behavioural phenomenon fostered at the individual level (e.g. McClelland 1961). More recently, 'psychologists have taken the lead in studying the social and cognitive development of children and adults – work that has had an important influence on how we understand the nature of change in the human development life-cycle' (Wagner 2018, 23–24).

Broadly speaking, three major learning theories – behaviourist, cognitive, and social constructivist – shape our perceptions of what counts as learning and which aspects of performance are valid for inference of competence (that is human capital, in economic terms). Behavioural learning theory emphasises meeting behavioural or instructional objectives (Watson 1930; Bandura 1969). Assessments under this approach use standardised tests in particular learning domains (e.g. mathematics or science), and they generally serve to evaluate the degree to which students have actually acquired the curriculum compared to what is presumed to have been taught. Cognitive learning theory draws attention to the development of children's cognitive abilities (i.e. intricate mental processes; Neisser 1967; Sternberg 1981). Assessment of cognitive abilities involves more than simply gauging knowledge and skills; it also appraises higher-level of processes, such as comprehension, application, analysis and evaluation (Frederiksen 1984; Sugrue 1995). In contrast, social constructivist theory conceives learning as socially constructed, and it underscores the dialectical relationship between the learner and the environment (Vygotsky 1978). The educational prosperity framework (examined

in Chapter 6), which posits that learning is the cumulative effect of a range of factors that affect children's development, can be located under this social constructivist approach.

The major approach that the OECD PISA draws on is cognitive psychology. In recent years, the OECD has been expanding the scope of PISA metrics to incorporate non-cognitive skills, such as social and emotional skills, creativity, and well-being (see e.g. Sellar and Lingard 2014; Grotlüschen 2018; Auld and Morris 2019; Rappleye et al. 2019; Li and Auld 2020). These non-cognitive dimensions are considered to 'represent more enduring *potential* for human capital appreciation, in contrast to specific technical or academic skills that are susceptible to obsolescence with technological and economic change' (Sellar and Lingard 2014, 926, original emphasis).

From the social constructivist perspective, however, the value of non-cognitive skills is strongly context-dependent, and their attributes can be 'irreducibly heterogeneous' (Bowles, Gintis and Osborne 2001). This means that to measure and compare these skills, the OECD subscribes to two major assumptions: first, that these skills are generalisable from specific contexts; and second, that they can be made commensurable cross-nationally based on common metrics. While these assumptions are identified by Morris (2015), the critical point here is that the 'translation' process is neither neutral nor merely technical. It is a way to reduce and simplify disparate attributes into numbers that can easily be compared. According to Espeland and Stevens (2008), commensuration transforms qualities into quantities and difference into magnitude.

So far, there has been limited work on the validation of constructs in educational assessments, partly because of the scope and complexity as discussed above and partly because the modern validity theory itself offers little guidance on how to perform this task in practice (see Wolming and Wikström 2010). It follows that 'the field is riddled with *assumptions* that the attributes test designers set out to assess *are* being assessed, and those attributes are themselves under-defined' (Baird et al. 2017, 322, original emphasis). Therefore, related to Carnoy and Rothstein's (2015) observation of ILSAs leading to inappropriate or even harmful educational reforms, Baird et al. (2017, ibid.) provide a possible answer: 'national policy-makers could hardly understand the content of the tests or the effects upon learning of changing them'.

In addition to the skill sets to be defined, ILSA design must address the two other key elements, population samples and contextual variables, in order to reliably estimate outcomes and meaningfully interpret them. There is no doubt that participants in any

assessment should be representative of the target population. However, when it comes to low- and middle-income countries, Wagner (2018) warns that poor and marginalised groups may be excluded from, or underrepresented in, samples that are supposed to be nationally representative. This often happens because of

> expediency (i.e. it may require too much time or too many resources to fully represent each socio-demographic stratum) and national politics (i.e. marginalised groups may not be considered important enough for policymakers and/or they may not be citizens). Thus, it is not unusual for national assessments to oversample the easier-to-reach urban areas, rather than rural areas. Furthermore, in some developing countries, the difficulty of (literally) tracking down nomadic children or studying children in fragile and conflict-ridden situations can make it difficult and expensive for education authorities to study such children. (Wagner 2018, 233)

The issue of representation in PISA-D is similarly controversial, especially Strand C which includes out-of-school youth. Although these concerns were raised in the project background working paper (see Carr-Hill 2015), in reality, PISA-D failed to address them (Auld, Li and Morris 2020). The long delay of the PISA-D report on out-of-school assessment results also indicates that PISA-D replicates rather than addresses these challenges widely discussed in the literature. However, considering that PISA-D results will be used for benchmarking the progress of the global education agenda, the degree to which these marginalised groups of children are included in the population samples has serious implications, which I discuss in Chapter 6.

ILSAs also develop contextual questionnaires for different participants (e.g. teachers, students, and administrators) to collect information for interpreting test results. As noted earlier, the process of analysing and interpreting data is critical for supranational organisations, such as the OECD, to shape national policy agendas (Martens and Jacobi 2010). Bloem (2015) also identifies that in order to increase the policy relevance of PISA, the OECD has shifted data analysis from external consultants to its own analysts. In the case of contextual questionnaires, the OECD has stressed the importance of the classroom and school factors for explaining student performance/achievement (Willms and Tramonte 2015).

However, the issue with placing too much emphasis on the 'school effects' is that it tends to overlook the more salient factors pertaining to processes and resources – another possible answer for understanding why the resulting analyses have provided misleading results. According to Willms and Tramonte (2015, 8), student performance in the PISA tests at age 15 'represents the *cumulative* result on children's experiences at home, in the community, and in school since birth, or even earlier' (original emphasis). As such, 'we

should not expect that measures of school or classroom practices, derived from questionnaires administered at the same time as the achievement tests, to have strong relationships with performance' (Willms 2018b, 7). Collecting contextual information is more crucial in low- and middle-income countries, because 'children's early family and school experiences likely play a more dominant role than in OECD member countries' (Willms and Tramonte 2015, 7).

In addition, ethnic/linguistic diversity is another important factor that matters significantly in these countries. In Africa, for example, Brock-Utne (2016) notes that both the content of assessments based on Western models and the language in which questions must be answered are 'foreign' to the majority of children. The annual reports from the Kenyan National Examination Council have consistently stated that participants failed to perform optimally in maths and science, 'either because they fail to understand the questions or because they cannot express themselves clearly in English' (Brock-Utne 2016, 39). The fact that the majority of children who do not speak the language of instruction (or test) at home further challenges the validity of ILSA results, especially PISA-D results.

Beyond an appraisal of technical issues

There are also concerns that extend beyond the appraisal of technical issues. Kaess (2018) highlights the epistemological difficulties for PISA-D in developing test material that would be culturally meaningful and internationally comparable. She warns of the risk that questions are 'ethnocentric', meaning 'inherently and intrinsically biased', and are, therefore, more difficult to discern and resolve:

> Some problems in the question material are perhaps more easily discernible – the inclusion of food items that are culturally unknown, or that are forbidden in specific cultural contexts, for instance – and can therefore be eliminated in the translations of the assessment. Epistemological differences, however, are potentially more deep-seated, less visible, and thus more difficult to bypass or counteract solely through acts of translation. (Kaess 2018, 352)

This warming reflects the more fundamental problem characterising PISA and PISA-D which Kaess (2018, 357) describes as 'knowledge colonialism'. That is, one system of knowledge is conceived of as superior to alternative systems of knowledge, and so the OECD promotes 'the existence of a single epistemological system valid for children all across the world' (ibid.). From the postcolonial perspective, the colonisation of language and thoughts, compared to the 'hard' forms (of extracting natural resources and depriving

economic independence), has a more profound effect on the 'otherness' of the colonised, by rendering them deficient and degenerate in relation to the norm (see e.g. Fanon 1967; Said 1980; Spivak 1988, 1999). For postcolonial scholars, the imposed language and frames of thought are means to legitimise the existing oppressive structuring of societies.

Extending colonialism beyond intercultural analyses, d'Agnese (2015) focuses on the internal logic of PISA. She suggests that 'colonialism is the very nature of PISA for reasons that are completely internal to PISA's philosophy' (ibid., 56). She identifies this logic operating through a hidden chain of PISA signification: linking education with learning, learning with assessment, and assessment with PISA. Under the guise of objectivity, she notes, PISA manifests a clear ideology, which places education in a well-defined value square: money, success, evidence, and competition. Therefore, for d'Agnese, PISA is not an assessment tool but an all-encompassing framework that 'expropriates culture and knowledge from subjects, denying their legitimacy and imposing the OECD's own univocal logic' (ibid.).

d'Agnese's point has some resonance with Tikly's (2004, 173) notion of a *new imperialism*, which incorporates populations within the formerly so-called 'Second' and 'Third worlds' into a regime of global government. PISA-D, in this regard, is a means of drawing low-income countries into the PISA international assessment regime that is already inhabited by many high- and middle-income countries (Tikly 2017). Resonating with Kaess and d'Agnese, this thesis considers the coloniality of PISA-D, and through which, as also identified by Tikly (2004, 2017), the OECD assumes the role of an emerging imperial power in global education.

Summary

McCowan (2015) notes that postcolonial analyses of educational applications are largely critiques that reveal the hidden assumptions and veiled disparagement of others' cultures and knowledge. Decolonising PISA-D, therefore, involves a process of acknowledging the multiple perspectives on reality and refraining from a dominant interpretation. This literature review has attempted to read from these different perspectives, broadly located within the fields of CE, IR, and EID. While there are some overlapping themes and critiques, the arrangement overall helps to better understand the multiple features that PISA-D embodies and the interrelated objectives which the OECD aims to achieve.

At a general level, the analyses that follow adopt the constructivist IR approach to interpret the origins of PISA-D and the evolution of OECD in global educational governance presented in Chapter 4; then, a managerial view of the OECD in legitimating PISA-D as a universal learning metric, and thereby establishing its role in monitoring SDG 4 in Chapter 5; and a critical examination of the outcomes of PISA-D in Chapter 6. All of these analyses are informed by postcolonial considerations attempting to reveal the hidden assumptions and modus operandi and to problematise and challenge the assumed authority and organisational legitimacy. The next chapter provides a methodological account of how these analyses are pursued through a critical hermeneutic approach.

Chapter 3. Research Methodology

Introduction

Kaplan (1964) highlights the difference between what is actually done in research and the way in which it articulates what is done. Specifically, he refers to the logic a researcher uses to produce knowledge as *logic-in-use*, and the attempts to explicitly formulate and articulate the logic-in-use as *reconstructions*. This chapter can be seen as an application of the reconstructed logic in that it seeks to describe the research position and process and also provides the explanation and justification of the research methods. In this regard, reconstructions are interchangeable with methodology and are useful for illuminating the logic-in-use. As Kaplan puts it, the aim of methodology is 'to describe and analyse […] methods, throwing light on their limitations and resources, clarifying their presuppositions and consequences'. Alternatively, the aim is to 'help us to *understand*, in the broadest possible terms, not the products of scientific inquiry but the process itself' (ibid., 11, original emphasis).

According to Punch and Oancea (2014), the research process can take one of two approaches: the first is paradigm-driven and the second is question-driven. The former focuses on understanding the assumptions and practices associated with particular research traditions (such as positivism or interpretivism) and deriving research questions and methods from them. By contrast, the latter begins with a problem or research questions that may come from any source (e.g. the literature, existing substantive theory, or practical issues or problems) and then chooses relevant methods to solve the problem or answer the research questions. Punch and Oancea also note that in higher degree research programmes at universities, the paradigm-driven approach is emphasised to the extent that research will not be allowed to proceed if it has not articulated its paradigm position. However, Punch and Oancea argue that these two approaches should be equally valued and the perception that all research must be paradigm-driven is no different from subscribing to dogma.

In this thesis, I have found that these two major approaches are not easily separable. While the question-driven approach may resonate more closely with the logic through which my research design emerged, there is also a strong engagement with philosophical issues. This is because the crafting of research questions reveals itself through the

paradigmatic traditions that it contains and rests upon. The research design is based on assumptions about the nature of the reality being studied, what constitutes knowledge of that reality, and what the appropriate ways are of generating knowledge of that reality.

In terms of the structure of this chapter, I begin with an account of the research focus based on observations and insights generated from the literature. These observations and insights are important to formulate the three main lines of inquiry and the corresponding research questions. I then present the research design with regard to how these questions are to be approached and answered. This is followed by a discussion on the three data generation methods (documents, interviews, and literature) and on how an intertextual reading of these sources could possibly be achieved to elicit meanings. Thereafter, I illustrate how this process of interpreting data would look like – a sort of 'upward moving mode'. In the last section, I reflect on the challenges I encountered and my ethical considerations during the study.

Developing the research focus: a puzzle from the literature

> Interesting work begins not just with a problem […] but with a puzzle. … Great leaps forward […] often take place when someone sees puzzles, where others have only seen facts.
>
> —Robert O. Keohane, *Political Science as a Vocation*

> We're an empire now, and when we act, we create our own reality, and while you are studying that reality … we'll act again, creating other new realities, which you can study too, and that's, how things sort out. We are history's actors.
>
> —Karl Rove, chief consultant and policy advisor to George W. Bush

As Chapter 1 stated, the OECD positions PISA-D as an important contribution to the monitoring and achievement of SDG 4. However, my readings of the literature in Chapter 2 suggest that the OECD's portrayal and empirical findings on the impact of PISA on development (both economic and educational) are rather conflicting. Moreover, to reiterate, PISA-D, as an exercise in CE, is an 'applied' form, attempting to measure and compare all types of countries with vastly different conditions and needs. The OECD's approach is premised on positivist and hyper-rational assumptions designed to provide policy solutions which contradict the 'academic' tradition that comparative studies should be context-driven and focused on understanding local cultures and histories. PISA-D, as it endeavours to serve as the precursor to a global learning metric, is based on shallow

philosophical and methodological grounds, and is oriented towards consolidating the status quo, including its injustices and inequalities.

This study begins to elicit the deeper meaning of PISA-D in the light of these controversies. The meaning can be fully grasped only by situating PISA-D within the organisation's long-term pattern of adaptation, survival and expansion. That is to say, PISA-D is seen here not as a singular event but as an integral part of the OECD's evolving strategy. This understanding resonates with the constructivist approach to IOs and their activities. Essentially, according to Bryman (2012, 34), constructivists 'consider the ways in which social reality is an ongoing accomplishment of social actors rather than something external to them and that totally constrains them'. Therefore, as researchers, 'we cannot take for granted, as the natural scientist does, the availability of a preconstituted world of phenomena for investigation' and must instead 'examine the processes by which the social world is constructed' (Walsh 1972, 19). In line with this constructivist paradigm, the study starts off by tracing the origins of PISA-D and illuminating the processes by which the contemporary form is constructed. Thus, it asks the corresponding research question: *what are the origins of PISA-D and how have they shaped its contemporary form?*

Having clarified the status of reality of PISA-D, the study proceeds to understand the OECD strategies for legitimating PISA-D as the universal measure of SDG 4. The second line of inquiry develops out of two related paradigmatic considerations. First, since the literature does not suggest positive relations between PISA and development, how does the OECD then ascribe the meaning of PISA to development? Second, the truth of such an attribution is similar to the source of organisational legitimacy which is constantly negotiated and skilfully managed. Indeed, studies of organisational theory show that legitimacy derives from a generalised perception or assumption of desirability, propriety and appropriateness, which organisations can manipulate to help achieve desirable outcomes or hidden agendas (Dowling and Pfeffer 1975; Suchman 1995). For example, 'organisations often put forth cynically self-serving claims of moral propriety and buttress these claims with hollow symbolic gestures', and some 'better integrated, more firmly established regimes tend to hold […] diverse legitimacy dynamics in alignment […] by defining certain arenas in which self-interest is considered morally laudable' (Suchman 1995, 579–585). From these perspectives, legitimacy management is a critical element in the process of the OECD construction. Therefore, the second line of inquiry aims to

answer the following question: *how has PISA-D been promoted and legitimated, and what strategies have been employed by the OECD to establish its role in monitoring SDG 4?*

The third line of inquiry attempts to deepen the preceding interpretations by examining PISA-D outcomes. Chapter 1 identified the lack of literature which systematically interrogates its outcomes (with the exception of Auld, Li and Morris 2020). The existing efforts have mainly remained at the level of 'arrival' and the level of the piloting process, avoiding any analysis of the end results of PISA-D, including the assessment instruments, findings, and policy recommendations. These outcomes are important to the thesis in that they encapsulate not only the past but also the future directions of the OECD and PISA-D. That is to say, the current outcomes carry with them the rationales and process as well as the possible future. As such, there is a need to provoke an earthquake-like examination, as Nóvoa and Yariv-Mashal (2003, 433) put it, 'in order to understand how these layers work, how they are connected and disconnected, producing contested explanations for the same "event"'.

Regarding PISA-D outcomes, there are two specific 'layers' I want to examine in this thesis. The first is related to the issues and challenges that low- and middle-income countries face in engaging in ILSAs (discussed in Chapter 2) and how PISA-D has addressed these challenges regarding relevance, comparability, and policy insights. The second elicits *specificities* of how the emerging OECD governance approach is operating, at the organisational and educational levels, and for the latter specifically, the connections between PISA-D and the existing PISA framework and standards as well as the new spaces for governance that PISA-D has just opened up. The corresponding research questions for the third line of inquiry are the following: *how has PISA-D addressed the challenges that low- and middle-income countries face in engaging in ILSAs, how are the assessment results generated and compared, and what policy insights do they provide for the piloting countries?*

An interpretive research design: positions, perspectives and processes

As stated above, the study is pursued through a critical interpretive research design. According to Schwartz-Shea and Yanow (2012), an interpretive research design is more dynamic and flexible, and focuses on the thought processes and ensuing strategies that go

into designing the study. The thought processes and strategies rest on a belief in the existence of multiple, intersubjectively constructed 'truths' about social reality and that the researcher's understandings are co-generated through interactions with those being studied. These premises reflect paradigmatic considerations at several levels, including the ontological belief that reality is socially constructed, the epistemological ground is based on multiple perspectives, and the researcher's position is part of the meaning-making process. As such, Schwartz-Shea and Yanow (2012, 2) encouragingly state that interpretivist researchers should not have their confidence undermined by uninformed critiques, such as comments that 'interpretive research does not stand on its own', or '[is] useful only as a preliminary stage to generate information that can serve as the basis for a quantitative study'.

Walsham (1993) notes that interpretive research should be judged according to how 'interesting' it is to the researcher as well as to those in the interpretive or epistemic community. For Taylor (1971, 5), '[a] successful interpretation is one which makes clear the meaning originally present in a confused, fragmentary, cloudy form'. However, how do we know that this interpretation is correct or clear? Taylor answers, '[p]resumably because it makes sense of the original text: what is strange, mystifying, puzzling, contradictory is no longer so, is accounted for'. These accounts reflect an appreciation of the centrality of meaning in human life and the requisite open-endedness and flexibility in interpretive research, points that I return to below.

Methodologically, an interpretive research design is underpinned by observation and interpretation, that is, collecting information about social actions and making meaning of that information with reference to the views of the social actors (Aikenhead 1997). It attempts to understand phenomena through the meanings that people assign to them (Deetz 1996). The interpretive research design thus finds its intellectual traditions in Max Weber's notion of *verstehen*, hermeneutics, phenomenology, and some critical theory (see Yanow 2009; Bryman 2012). In order to understand these traditions and how each of them impacts on my own thinking of the research design, a brief account is helpful.

In German, *Verstehen* means understanding; it was used by Weber (1947, 88) to refer to his approach to sociology, 'a science which attempts the interpretive understanding of social action in order to arrive at a causal explanation of its course and effects'. While the Weberian approach seems to embrace both explanation *and* understanding, it maintains that the task of 'causal explanation' be undertaken by drawing on 'interpretivist

understanding of social action' rather than on external forces that have no meaning for those involved in the social action (Bryman 2012).

Hermeneutics originates from theology focusing on interpreting Biblical texts. When applied to the social sciences, it was extended to the interpretation of various human expressions embodied in conversations, texts, and physical artefacts. Hermeneutic thinkers contribute to articulating a set of guidelines which can be agreed upon within an interpretive community for making sense of those expressions (e.g. Ricoeur 1971; Taylor 1971). The central idea behind hermeneutics, according to Bryman (2012), is to bring out the meanings of texts from the perspective of their authors, entailing sensibility to the social and historical context within which these texts were produced. This idea is also shared by phenomenological studies as shown below. Schwartz-Shea and Yanow (2012) tend to emphasise the 'meaning-focused, semiotic character' of hermeneutics, and observe its wider application:

> Ethnographic, participant observation, ethnomethodological, and other modes of analysis are [all] infused with this understanding of their meaning-focused, semiotic character: they seek to elicit meaning by rendering spoken words and/or acts, and the objects referred to or used in these, as written texts and applying to them a hermeneutic analytic sensibility. (ibid., 42)

Phenomenology focuses on the common-sense thinking of individuals. It concerns how individuals make sense of the world around them and how, in particular, the researcher should 'bracket out' his or her preconceptions in order to grasp the social reality. This is because, as Schutz (1962, 59) argues, 'the observational field of the social scientist – social reality – has specific meaning and relevance structure for the beings living, acting, and thinking within it'; therefore, in order to grasp this social reality, the thought objects constructed by the social scientist have to be based upon the thought objects constructed by the common-sense thinking of those being studied. The implication for this thesis is that the meaning of PISA-D can only be elicited by approaching it from the OECD's 'thought' processes; data generation will involve a dialogical engagement with sources that are pertinent to the OECD (see the section on data generation below).

A related consideration in this respect is Kaplan's (1964) point about 'process' that characterises social science research. Similarly, Becker (1998, 61) suggests that if we want to understand how a phenomenon or event occurred, do not just look for the factors in the background or the present circumstances; instead, look for the story of how it occurred, 'for all the steps in that processes, for all the steps connected to each other, for

how one step created the conditions for the next step to occur'. The implication for understanding PISA-D is thus to trace its origins and to identify the processes by which they have shaped its contemporary form. This way of approaching PISA-D echoes von Wright's (1971, 57–58) notion of retrodiction: 'From the fact that a phenomenon is known to have occurred, we can infer back in time that its antecedent necessary conditions must also have occurred, in the past. And by "looking into the past" we may find traces of them (in the present)'. I suggest that, in addition to the past in the present, there are also conditions already created for the future. Accordingly, combining these considerations, Chapter 4 provides a periodisation of the OECD's changing agendas and approaches to education (from the 1950s to 2030), through which the origins, requisite conditions and significance of PISA-D can be better understood. To provide a general idea, these distinctive periods are schematically represented by the circles in Figure 3.1.

This is not to suggest, however, that my interpretations are uncritical or neutral. Critical theory also has its intellectual influences in two aspects: first, critical theorists believe that there is no value-free or theory-free natural and social scientific inquiry; second, critical theorists argue for a strong sense of value involvement. The latter aspect can be read from critical theorists' criticism of phenomenologists for being too preoccupied with individual meaning-making at the expense of considering institutional constraints, including power (Schwartz-Shea and Yanow 2012). Bryman (2012) concurs, noting that social scientists will almost certainly aim to place their interpretations into a social scientific frame. As such, there are two layers of interpretations: 'the researcher is providing an interpretation of others' interpretations', and 'the researcher's interpretations are further interpreted with reference to the concepts, theories, and literature of a discipline' (ibid., 31).

In this thesis, I have drawn on constructivism (from IR discussed in Chapter 2) as the broader theoretical frame to interpret the OECD's evolving role in global educational governance and the processes by which PISA-D was concretised into a core element in sustaining and extending the organisation's influence globally. I have also embraced an instrumental/managerial view of organisational legitimacy to illuminate the OECD's strategies to legitimate PISA-D and establish a clear role for itself in monitoring SDG 4. Moreover, the preceding interpretations and those of the PISA-D outcomes (including instruments, findings, and policy recommendations) have been informed by postcolonial scholarship concerning the extension of ILSAs into low- and middle-income countries

(as an act of colonialism or new imperialism). As such, the design of this study itself reflects an effort to 'decolonise', entailing 'a process of "learning to read the world through other eyes", of acknowledging the multiple perspectives on reality and reframing from a single reading' (McCowan 2015, 43). Figure 3.1 sketches out how these perspectives and concepts are applied in relation to the corresponding research questions.

Clearly my interpretations pay critical attention to power dynamics as well as silent, and silenced, discourses. I thus describe my research method as endorsing a *critical hermeneutic* approach (e.g. Habermas 1974). Similar applications of this approach can be found in Phillips and Brown (1993) and Forster (1994). These researchers specifically focus on organisational communication through interpreting company documents. Phillips and Brown (1993, 1547), for example, note that the critical sense of their approach lies in 'self-conscious reflection on the social conditions surrounding the production, dissemination, and reception of texts and on their contribution to the creation and maintenance of power differentials in and around organisations'. Similarly, Forster (1994) interrogates company documents and extracts themes from them by reference to his knowledge of the organisational context in which the documents and the people and events within them were located. These applications not only allow an understanding of conditions and processes to be forged, but they also offer further insights into the treatment of documents as a major source of interpretive data. In the following sections, I elaborate on my data generation methods and my adaptation of the hermeneutic process that Forster (1994) developed to interpret the OECD documents.

After describing the major approach and perspectives, here I add some reflexive thoughts on the research process which are no less important for this research design. Initially, the literature review mainly contained two bodies of literature drawn from the fields of CE and EID (i.e. key themes and perspectives summarised in the left and right boxes in Figure 2.1). However, after I encountered the 'field' realities, I realised that they were inadequate for understanding the meaning of PISA-D, especially the multiple purposes that it serves to achieve, and which are intertwined with the OECD strategy for development and the UN-led SDGs. These realities indicate a more focused scholarship on (i) the changing role of the OECD as a purposive actor in global educational governance, and (ii) the major sources of power through which it exerts influence. In response, I then added two further bodies of literature broadly drawn from IR scholarship and global governance research to incorporate these themes. As shown in Chapter 2 (at

the bottom of Figure 2.1), I accomplished this by referring to existing IR theories and the operating mechanisms of OECD global governance in education, as well as ideas about organisational legitimacy and the politics of expertise. Accordingly, with these new insights and better 'situated' knowledge, I amended the research questions for greater specificity and proceeded with further interpretations. This type of process was repeated several times until the last phase of the completed manuscript.

These steps and considerations illustrate that an interpretive research design is not a set of fixed, standard arrangements, but an iterative and reflexive *process* of thinking back and forth. The process is described by Schwartz-Shea and Yanow (2012, 56) as follows:

> The rhythms of interpretive research are quite different. Its design revolves around front-loading of a different sort: the researcher's provisional sense-making will be "tested" in the field or archives, not literally, as in the case of significance tests in statistics, but by bringing them together with field realities, even to the point where those realities might take analytic primacy. Initial understandings are likely to be reformulated in light of new insights, new understandings, new knowledge acquired, and those reformulations will be subjected again to further inquiry, in that iterative, spiral-circular recursiveness of abductive reasoning […]. It is this continuous juxtaposition of conceptual formulations with field realities and the requisite flexibility that accompanies it that compose the fundamental rhythm of interpretive research.

I would also add that it is these features of openness and flexibility and the researcher's willingness to reflect on the recursive processes that shed light on the underlying logic of interpretive inquiry, on what Kaplan (1964) refers to as the *logic in use*. In Figure 3.1, it is only in retrospect that these key elements and/or parts can be composed as a whole and presented in a relatively patterned manner.

Figure 3.1 Sketching out the interpretive research design on PISA-D – *reconstructed logic*.
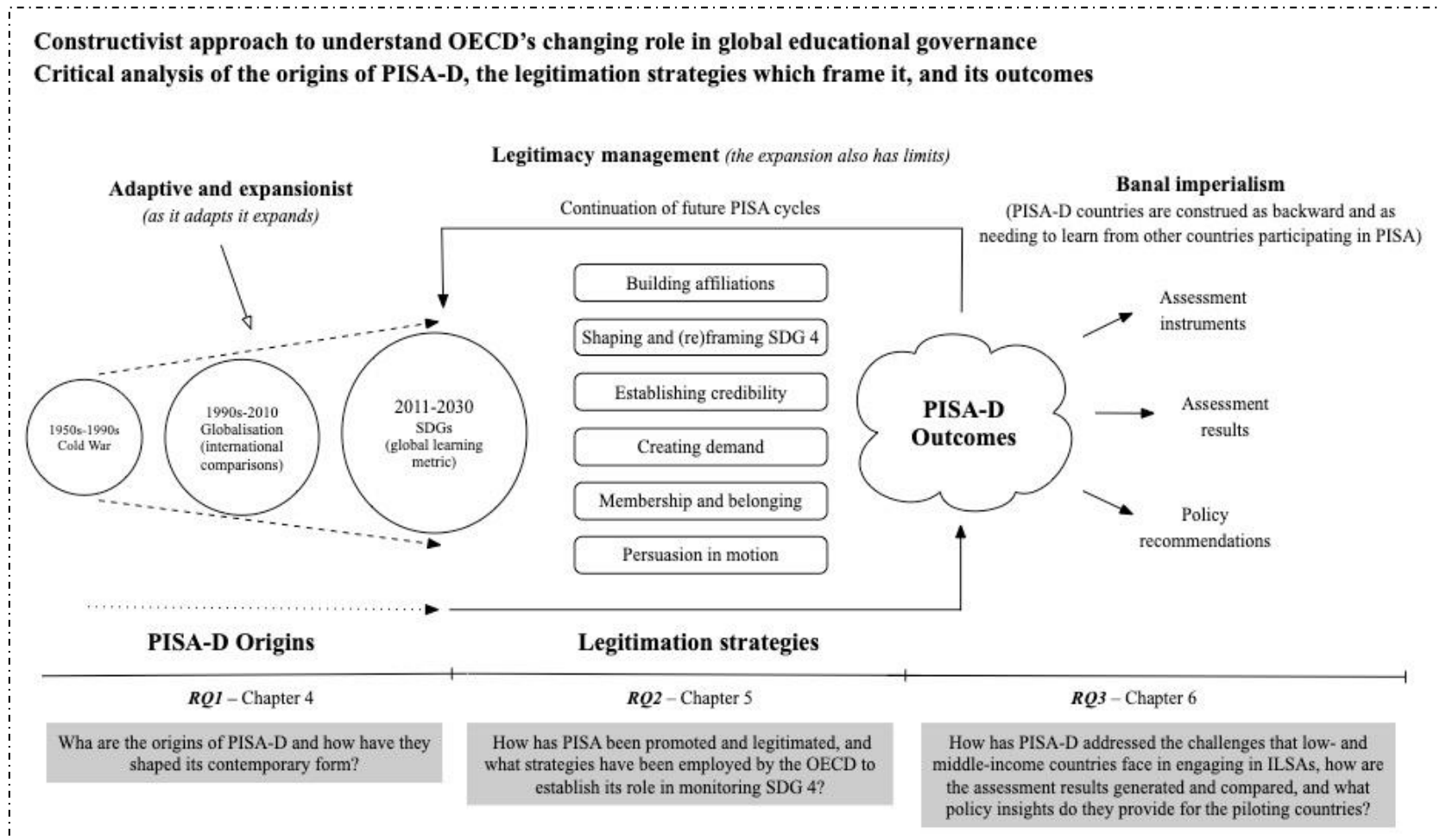
Data generation methods and intertextuality

The preceding section (and Figure 3.1) outlined the major research approach and strategies for this study. OECD documents and semi-structured interviews served as the primary source of information. Secondary literature was another important source. The nature of interpretive research suggests that the distinction between data collection and data analysis is not clear-cut. It may be helpful to consider them as 'an organic whole' in which interpretation 'begins in the data-gathering stage and does not end until the writing is complete' (Potter 1996, 120). Therefore, I prefer to use the term 'data generation' instead of data collection. Below, I provide a detailed account of the methods and discuss the possibility of 'reading across' these different types of data.

Documents

Documents are an established and appealing source of data in interpretive research. Like other analytical methods, documentary analysis scrutinises data to elicit meaning, gain understanding and develop situated knowledge (Corbin and Strauss 2008). Atkinson and Coffey (2004) consider documents as 'realities' that are produced, shared and used in socially organised ways. They note that in addition to using documents for record-keeping, organisations also produce significant documents of various kinds that are concerned with their *self-representation*. These accounts relate to Phillips and Brown's (1993) point about the communicative purpose of organisational documents, which is not merely informative but primarily transformative by attempting to link ideas and symbols in ways that contribute to the creation or maintenance of an enduring pattern of social relations. As such, Phillips and Brown suggest that '[a]ny activity or object that affects the network of symbols through which a group or individual understands an organisation and its place in the world is of interest from a critical hermeneutic perspective' (ibid., 1548).

Of course, in contemporary times, as Coffey (2013) points out, we should also include electronic and digital resources among the ways in which documentary realities are produced and consumed. These include, for example, websites, promotional videos, networking sites, and similar artefacts. These documents can be seen as the 'physical traces' of social settings (Webb et al., 2000) in which social actors act and interact. They are all among the techniques and resources that organisations deploy to create versions of reality and self-representations (Atkinson and Coffey 2004). Foster (1994) argues that

these varied documentary records are more comprehensive in coverage than the kind of material that a researcher gathers from interviews or questionnaires. In time, they often become contemporaneous records of events, helping the researcher to look more closely at historical processes and developments in organisations. This knowledge can, in turn, add depth to the questions asked during interviews and assist in interpreting interviewees' 'rewriting' of history in verbal accounts.

In the case of the OECD, documents relating to PISA-D are mainly located on the PISA-D webpage[2] composed of four major sections: (i) *About* (including Background and one Brochure); (ii) *Documentation* which includes Publications (i.e. PISA-D Technical Reports, PISA-D Assessment and Analytical Framework, Review of ILSAs in Education, and the Experience of Middle-income countries in PISA 2000–2015), Country Reports (including Capacity Needs Analysis, Capacity Building Plan, and national reports), Working Papers (on the cognitive instruments, contextual questionnaires, and developing the out-of-school survey), Blogs, and PISA-D Briefs; (iii) *Country Participation* (i.e. information about the responsible institution(s) within each piloting country and relevant media coverage); and (iv) *PISA-D Results*, which include the PISA-D Results in Focus, PISA-D databases, and national reports. There is also a short video documenting how the OECD's education work is related to the education SDG.

While these documents provided on the webpage are extensive and almost all-encompassing, there are two additional significant sources of documents that are stored on the OECD website but not on the PISA-D page. The first covers the PISA-D meetings and workshops held globally from 2013 to 2017[3], and hence all PowerPoint presentations that were given; the second is the PISA-D International Seminar held in London in 2019[4] where the PISA-D reports were launched, outcomes and findings were presented, and the success of PISA-D was declared. These meetings, workshops, and seminar, along with the large volume of presentations, recordings, and summary report, provide rich documentary records for tracing and examining PISA-D processes and outcomes. They also reflect the social settings in which these documents were produced and circulated. By interrogating these data, they allow me to move beyond the surface issues of sense-

---

[2] https://www.oecd.org/pisa/pisa-for-development/

[3] https://www.oecd.org/pisa/aboutpisa/pisa-for-development-meetings.htm

[4] http://www.oecd.org/pisa/pisa-for-development/pisafordevelopmentinternationalseminar.htm

making to examine the means by which the OECD came to legitimate its position and the activities of PISA-D.

Atkinson and Coffey (2011) argue that while it is tempting for the social scientist to regard documentary materials as secondary data – to help cross-check observational or oral accounts, or to provide some descriptive and historical context – documents should be treated as data in their own right. They should be given due weight and appropriate analytical attention. In this thesis, I have considered the rich set of PISA-D documents as an asset and have treated them as the primary source for identifying the legitimation strategies and examining PISA-D outcomes (see Table 3.1 below for a list of documents). I have also drawn on interview data and the literature to help pursue these goals, including tracing the origins of PISA-D and seeking out multiple possible views (that may be inconsistent or even contradictory) on PISA-D. I elaborate these additional sources in the sections below. Although this practice of using multiple sources of data embodies triangulation for recognising the complementarity of different methods, it actually moves beyond triangulation in considering the *dialectical* interdependence of different sources and the researcher's upward-moving mode of understanding.

For example, before interviewing the OECD staff, I had not known much about the overarching framework (i.e. DeSeCo) underpinning all of the OECD's assessment programmes, including PISA and PISA-D (see Chapter 4). This knowledge pointed me towards locating the relevant documents which in turn have greatly facilitated my understanding of PISA-D, not as a singular new project, but as deeply rooted in the OECD assessment regime whose origins can be traced back to Cold War politics. The finding that PISA-D is a self-serving instrument, however, conflicts with the OECD's proclamation that it is a development project tailored to the needs of low- and middle-income countries. By examining PISA-D reports and seminar documents (focusing on the results, modes of comparison, and policy recommendations), I came to confirm my earlier finding and then diverted attention to the literature that perceives IOs as autonomous sources of power with their own interests. These are just two examples illustrating how the different sources of data I have drawn on 'talk' to each other and provide a higher-order level of understanding (also see Figure 3.3 for the hermeneutic process of data interpretation).

In this respect, 'intertextuality' is a more suitable term which implies that the researcher 'reads across' different types of data in eliciting meaning and gaining a deeper

understanding. Schwartz-Shea and Yanow (2012) suggest that intertextuality is a marker of research quality in interpretive studies, because it means taking participants' views seriously and embedding their meaning-making in its social, political, cultural, and historical contexts:

> Intertextual readings of this sort look for the dimensionality, ambiguity, and possible contradictions that might arise from broad examination of evidence, the researcher remaining open to the possibility of consensus and agreement without presuming or privileging it. It is seeing this intertextuality, and drawing on it in analysis, that leads to the "thickness" of interpretation […] in ways that enable analytical sense-making. (ibid., 86)

Table 3.1 Main documents featured.

Research questions		Thematic clusters	Titles and publication dates	Intertextuality
RQ1. What are the origins of PISA-D and how have they shaped its contemporary form?	Chapter 4	OEEC, productivity Economic recovery Scientific knowledge Technical training	The U.S. FOA Productivity Programme Memo (1954) The U.S. National Defense Education Act (1958)	Secondary literature (see Table 3.4)
		Executive Committee Member countries Knowledge-based economy Global competitiveness Standards, targeted learning measures Manage knowledge Central steering International INES, CERI DeSeCo PISA	*The Acts of the Organisation (Vol. I)* (1990) *Meeting of the Education Committee at Ministerial Level* (1990) *Education at a Glance* (1992) *Governance in Transition: Public Management Reforms in OECD Countries* (1995) *Shaping the 21st Century: The Contribution of Development Co-operation* (1996a) *The Knowledge-Based Economy* (1996b) Background Paper – *Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo)* (2001) Strategy Paper – *Definition and Selection of Competencies (DeSeCo): Theoretical and Conceptual Foundations* (2002) *Summary of the Final Report "Key Competencies for a Successful Life and a Well-Functioning Society"* (2003b) *Innovation in Knowledge Economy: Implications for Education and Training* (2004) Executive Summary – *The Definition and Selection of Competencies* (2005b) *OECD's Centre for Educational Research and Innovation—1968 to 2008* (2008) *Historical Summary of CERI's Main Activities* (2012c)	Interview OECD staff (see Table 3.3)
		Anniversary Strategy on development	*Better Policies for Development: Report on the DevGoals Exercise* (2011a) *Framework for an OECD Strategy on Development* (2011b) *OECD 50th Anniversary Vision Statement* (2011c)	

		Comprehensive approach	*The High-level Fora on Aid Effectiveness* (2003a, 2005a, 2008, 2011b)	
		Post-2011 governance	*Secretary-General's Strategic Orientations for 2011 and beyond* (2011d)	
		SDGs	*Background Report for the OECD Strategy on Development* (2012a)	
		Aid effectiveness	*OECD Strategy on Development* (2012d)	
		Developing countries	*Progress Report on the Implementation of the OECD Strategy on Development* (2013b)	
		PISA, PISA-D	*2014 Report on the Implementation of the OECD Strategy on Development* (2014a)	
		Next steps	*Looking Ahead to Global Development Beyond 2015: Lessons Learnt from the Initial Implementation Phase of the OECD Strategy on Development* (2015)	
		Configuration	*Bridging the Data Divide for Development* (Angel Gurría 2017)	
		Differentiation	*Development Co-operation Report 2017: Data for Development* (2017a)	
		Data for development	*PISA for Development Project Completion Report* (2019c)	
		Data revolution		
RQ2. How has PISA-D been promoted and legitimated, and what strategies have been employed by the OECD to establish its role in monitoring SDG 4?	Chapter 5	Post-2015 agenda Education quality UNESCO UNICEF OECD contributions Experience, expertise PISA, policy analysis international collaboration Learning outcomes Secondary education Indicator 4.1.1 Measuring quality Learning metric Global scale PISA-D	Andreas Schleicher Presentation at the Initial Technical Meeting (2013a) *The OECD's Post-2015 Reflection Series* (2013c) – *The OECD's Contribution on Education to the Post-2015 Framework: PISA for Development* (including an audio recording by Michael Ward) Brochure – *Improving Learning Outcomes Worldwide: How PISA Can Help* (2013d) *PISA for Development Project Document* (2013e) Blog – *PISA for Development and the Post-2015 Agenda* (Davidson, Ward and Palma, OECD 2014) Blog – *PISA for Development* (Solheim, OECD 2015) Blog – *The Challenges of Widening Participation in PISA* (Schleicher [OECD] and Costin [World Bank] 2015) *Making Education Count for Development: Data Collection and Availability in Six PISA for Development Countries* (2016b) Brief 2 – *PISA for Development: Benefits for Participating Countries* (2016e) Brief 4 – How is *PISA for Development Being Implemented? (2016f)* Brief 27 – *Peer-to-Peer Learning to Strengthen Dissemination of PISA for Development Results* (2018f)	Interview OECD staff & informant

	Leading reference Comprehensive, rigorous Premier yardstick Country examples Brazil, Peru, Vietnam Emerging demand In response to Global PISA community Capacity building Peer-learning	UIS GEM Report Blog – *SDG 4 is a Universal Agenda – and That Includes High Income Countries* (2016) Michael Ward Presentation at the Capacity Development Workshop: PISA for Development (2016) Michael Ward Opening Remarks at the 5th National Project Managers Meeting (2017) *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science* (2018b) *Summary Record* of *PISA for Development International Seminar* (2019d) Silvia Montoya Presentation at the PISA for Development International Seminar (2019)	Secondary literature
RQ3. How has PISA-D addressed the challenges that low- and middle-income countries face in engaging in ILSAs, how are the assessment results generated and compared, and what policy insights do they provide for the piloting countries?	Chapter 6	Pilot project Accessibility, relevance Cognitive instruments Contextual questionnaires Out-of-school Grade repetition Outcomes	*Call for Tender 100000990 - PISA for Development Strand A and Strand B* (2014b) Working Paper – *PISA for Development Technical Strand C: Incorporating Out-of-School 15-Year-Olds in the Assessment* (Carr-Hill 2015) Working Paper – *Towards the Development of Contextual Questionnaires for the PISA for Development Study* (Willms and Tramonte 2015) *Capacity Building Plan: Cambodia* (2016a) Brief 15 – *Sampling and Survey Operations for the PISA-D Out-of-School Assessment Field Trial* (2017b) Francesco Avvisati Presentation at the PISA-D International Seminar – *Enhancing Cognitive Instruments for PISA-D* (2019) Ann Kennedy Presentation at the PISA-D International Seminar – *Enhancement of Cognitive Instruments for PISA-D* (2019) Andreas Schleicher Presentation of *Successful Schools in Testing Times: Insights from PISA 2018 Volume V* (2020) *PISA for Development Out-of-School-Assessment Results in Focus* (2020)	Interview OECD staff
		Rationales PISA-D findings Prosperity outcomes	*PISA for Development Results in Focus* (2018c) *Education in Cambodia: Findings from Cambodia's Experience in PISA for Development* (MoEYE 2018) *Education in Ecuador: PISA for Development Results*	

|		
Attainment, achievement	(National Institute for Education Evaluation 2018)	Interview technical expert
Foundations for success	*Guatemala in PISA-D: Programme for International Student Assessment* (MoE 2018)	
Resources	*Education in Paraguay: Findings from Paraguay's Experience in PISA for Development* (MoES 2018)	
School and community support	*Education in Senegal: PISA for Development 2017 Results* (MoE 2018)	Secondary literature
Policy options	*Education in Zambia: Findings from Zambia's Experience in PISA for Development* (Examination Council of Zambia 2018)	
Low hanging fruits	Brief 25 – *Strengthening Analysis and Reporting in PISA for Development Countries* (2018e)	
Quick wins, must haves		
Long-term	UIS Report – *Learning Divides: Using Data to Inform Educational Policy* (2018)	
	Brief 29 – *Meeting Expectations? Reflections on PISA for Development from Participating Countries* (2019)	

Interviews

Interviews are used extensively in qualitative and interpretive research due to their flexibility and relative efficiency (Bryman 2012). As an implicit phenomenological and hermeneutic mode of understanding, they are applied to understand a phenomenon from the perspective of the interviewees. As Kvale (1983, 174) rightly puts it, the interview 'makes it possible for the interviewees to organise their own descriptions, emphasising what they themselves find important'. To achieve this goal, the format of the interview is usually 'semi-structured' – it is neither a free conversation nor a highly structured questionnaire. The latter is typically applied in quantitative research.

In interpretive research, interviewees are participants who actively shape the course of the interview rather than passively responding to pre-set questions (King 1994; Walby 2010; Schwartz-Shea and Yanow 2012). Moreover, the relatively unstructured nature of semi-structured interviews allows the interviewer to probe more intensively into the meaning of the described phenomenon, especially when there are new aspects or insights instigated by the interviewee. As such, the process is both interactive and interdependent. Sullivan (1954) conceives the interview as an 'interpersonal situation' in which the relevant data are constituted by the interaction itself in the specific interview situation co-created by both sides. In participant observation, he notes, it is the interviewer, as a person, who is the method, the instrument.

In this study, semi-structured interviews were the preferred approach; they were carried out with key actors involved in PISA-D, including two OECD staff members, one technical expert, and one lead analyst from PISA-D participating countries. The OECD positions PISA-D as a highly collaborative project whose implementation consisted of five phases (see Table 3.2 reproduced below). With regard to the implementation schedule, one of my interview participants was an advisor involved in the out-of-school component in Phase 1 on the initial design; another interview participant was involved in Phase 4 as a lead analyst interpreting the results and writing the national report. The two participants from the OECD were responsible for the overall management of PISA-D, and hence were knowledgeable about all the other phases, including technical development, field trials and in-country data collection processes, as well as post-pilot governance. Overall, these interviews, along with the documents gathered from PISA-D meetings and training events, served as important sources for gaining an insight into the implementation of PISA-D and for addressing the research questions.

Table 3.2 PISA-D implementation schedule.
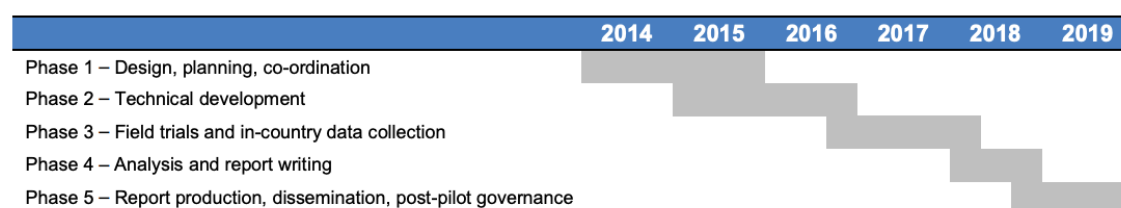
Table 3.3 PISA-D implementation: interview participants.

Themes co-generated during interviews	Participants and method(s)
Out-of-school component – Role and position; possible drivers for PISA-D countries; idea/logic of testing out-of-school children; sources of funding; tender process; contractors; technical issues and design of out-of-school survey Phases 1 & 2 & 3	Technical expert (1) Participant located in France (i) date: 07/08/2018 (ii) questions first answered in written form and sent back (iii) two follow-up emails
OECD perspective(s) – 1. Origin of PISA-D; rationales; suitability of PISA for low-income nations; PISA-D and SDG 4; OECD relations with UNESCO; OECD paradigm and UNESCO paradigm. post-pilot governance; OECD Education 2030; the DeSeCo 2. Cognitive instruments; contextual questionnaires; out-of-school survey implementation; selection of countries; sources of funding; communication; field trials; data collection, results, and post-pilot governance Phases 1-5	OECD staff (2) 1st participant located at OECD office (i) date: 17/10/2018 (ii) by phone; 1hour 2nd participant located in London (i) date: 23/10/2018 (ii) face-to-face; 40 minutes
Participating countries' perspective(s) – 1. Role; rationales; past experience in participating in PISA; working with OECD and technical partners; challenges; initial findings from out-of-school survey; report production; next steps for assessment activities in the country; use of findings in local context 2. Role; arrival of PISA-D in Cambodia; World Bank and OECD; internal discussions; resistance from UNESCO and UNICEF regional staff; key local actors' involvement 3. Capacity/receptivity of education officials in Paraguay; PISA-D; challenges; motivations; political conflict; capacity building; field trials; expected results Phases 1 & 3 & 4	Lead analyst (1) (i) date: 23/09/2019 (ii) by phone; 1.5 hours Informants (2) 1st informant located in Kyoto (i) date: 09/02/2018 (ii) face-to-face; 2 hours 2nd informant located in London (i) date: 21/09/2018 (ii) face-to-face, 1 hour

As Table 3.3 shows, I conducted the interviews through multiple formats: face to face, by email, and telephone. I had initially approached a wider range of potential participants to improve the likelihood of gaining access. However, there were unexpected and challenging experiences with approaching participants, particularly those in PISA-D countries. In some situations, after several rounds of communication (through email and contacts), even after the stage of gaining approval for interviews, the potential interviewees withdrew without explanation or further notice. This happened, for example, during my communication with representatives in Ecuador and Guatemala. In other circumstances, I received no reply after sending out email invitations. A similar experience also occurred when I wrote to the National Project Manager (NPM) in Honduras for a copy of their national report on PISA-D results which had not yet been made available on the OECD website. The limited access to potential interview participants was then accounted for in the research design, leading to an adjustment of the research questions and the hermeneutic interpretive approach. Addey (2016b), who has previously worked at UNESCO, acknowledges the difficulty in gaining access while conducting the PISA-D for Policy (P4D4Policy) project, noting in her methodology that the semi-structured interviews are the result of a long process of negotiating access and building trust, which took place over four years.

Schwartz-Shea and Yanow (2012, 71–73) contend that a researcher who expects her research to be implemented *as designed* is likely to be frustrated:

> Interpretive research designs *must* be flexible due to field realities, stemming from participants' agency. The researcher lacks control over research participants as well as over unfolding events… Changes in design cannot, therefore, be understood as threats to the trustworthiness of a research endeavor. (original emphasis)

According to Schwartz-Shea and Yanow, the very lack of control means that interpretive researchers cultivate particular competencies and skills, which are not required of positivist researchers, to manoeuver effectively in the field(s). The first is to *accept* the limited control, recognising that research is humanly and dynamically relational in form and character. The second is to *adapt to* the field realities, being willing to revise thinking in the light of new experiences and also being tolerant of ambiguity and uncertainty. The third is to *improvise* skills and the ability to manage the hybrid details of documents, interviews, and literature encoded in words and to apply appropriate analytical techniques. Improvisation does not mean 'making things up'; as Schwartz-Shea and Yanow argue, 'improvisation is a skill that is learned and perfected through practice'

(ibid., 76). It demands developing a methodological stance that is 'poised for movement, sensitive to flows and trajectories, to instability and reconstruction, to disruptions and re-orientations' (Zabusky 2007, 378).

A preference for flexibility and improvisation also applied to my interview process. For each interview, I had pre-listed questions that helped address my research questions but did not slavishly follow them in the actual 'interpersonal situations'. The primary focus was on how my interviewees described and made sense of PISA-D, including the origins, rationales, and details relating to the different implementation phases (see Table 3.3 for the themes co-generated from the interview process). For instance, during my interviews with the OECD staff, the relatively open questions allowed me to obtain accounts of how PISA-D (while officially presented as a development project) is deeply connected to the OECD's assessment regime, aiming for its global expansion and the requisite legitimacy (see Chapters 4 and 5). Improvisation was also important given the status and social power that my interviewees possessed, as members of bureaucratic institutions and experts in their areas. Similarly, Bogner, Littig and Menz (2009) also point out the less equal dynamics in the political dimensions of interviewing experts. In this regard, sound preparations, knowledgeability (e.g. having gone through the documents written about and/or by my interview participants), emotional sensibility, and self-confidence helped decrease the status imbalance between the interviewer and her interviewees (Zuckerman 1972; Berry 2002; Harvey 2011; Thuesen 2011; Mikecz 2012).

Literature

In this study, the literature served as another important source of data. While secondary literature presents findings that are processed by their original authors, they can serve as potential ingredients for pursing a new research strand. This is because secondary analysis is not merely descriptive but constructive. This idea is shared by Maxwell (2009, 223): 'It incorporates pieces that are borrowed from elsewhere, but the structure, the overall coherence, is something that *you* build, not something that exists ready-made' (original emphasis). Becker (2007, 142) likens the use of literature to doing a woodworking project:

> Other people have worked on your problem or problems related to it and have made some of the pieces you need. You just have to fit them in where they belong. Like the woodworker, you leave space, when you make your portion of the argument, for the other parts you know you can get. You do that, that is, *if* you know that they are there to use. (original emphasis)

The analogy from Becker points to the strategic use of literature (e.g. as 'modules') in constructing arguments in my own study. It also stresses the intrinsic value of literature, as an essential component of research rather than just the basis for it.

In this study, as the list in Table 3.4 shows, the literatures were selected and used based on their potential contribution to understanding the educational activities of the OECD (along with those of its predecessor – the Organisation for European Economic Co-operation [OEEC]) in the early decades, and how they have evolved into the contemporary forms, as well as where they are leading to in the near future. In other words, these are important pieces to incorporate while tracing the origins of PISA-D and identifying the ideological bases of the OECD's changing modes of governance in education. Moreover, many of these studies are authoritative sources of information. For example, I have drawn on the work of Kjell Eide (1990) and George Papadopoulos' (1994) because they both had played key roles in the development of the OECD's educational activities from the early 1960s. Professionally, Kjell Eide, a Norwegian economist, was a key figure in mapping the scientific and technical personnel, studying the economics of education, and carrying out educational planning at the OECD. George Papadopoulos was the founding deputy director for education at the OECD Directorate for Social Affairs, Management and Education for 20 years. Their accounts are rare sources of historical information for understanding how the OECD became involved in education and how its involvement developed overtime from the 1960s to the 1990s.

Another strand of literature pays attention to the global geopolitical and economic realities which have shaped the OECD's priorities and modes of operation. I used this literature particularly for illuminating the OECD's highly evolved capacity to both respond to global changes and play an integral part in those processes. Thus, for understanding how the OECD's educational activities relate to Cold War politics, I drew on the work of Mundy (1998, 2007) and Tröhler's (2013), and for the OECD and globalisation, I consulted Henry et al. (2001) and Rizvi and Lingard (2009). With regard to how the OECD's approach to education had shifted during these two periods (from the Cold War to globalisation), I delved into Martens' (2007) examination of the OECD's 'comparative turn' (thereby the experiment of PISA), followed by Sellar and Lingard's (2014) account of the OECD's expansion through the evolution of PISA. Following this thread, I then incorporated Addey's (2017), Auld, Morris and Rappleye's (2019), along with Moss and Urban's (2019) findings to capture the most recent developments at the

OECD, including PISA-D and its future implications. Moreover, to elicit the historical roots of these contemporary manifestations, I further built on Bürgi's (2019) work on the activities of the European Productivity Agency (EPA)/OEEC and Elfert's (2019) on the 'economics of education' approach.

Data from two additional interviews undertaken by others were also pertinent: one was Addey's (2016a) interview of Andreas Schleicher on PISA-D, and the other was a UNESCO IIEP[5] (2016) staff member's interview of Michael Ward on PISA-D. Taken together, the literature accounts for an important source of data for specifically addressing the research questions about the origins of PISA-D and the processes by which they have shaped its contemporary form (see Table 3.1 for the intertextuality of different sources). In doing so, it has served two purposes: the first was to understand PISA-D from the OECD's point of view by situating it within an analysis of the organisation's long-term pattern of adaptation, survival and expansion; and the second was to illuminate the OECD's changing role in global educational governance – having evolved from an instrument of the U.S to a supranational organisation with its own interests and enormous power. The central thread running through the changing trajectory (or the overall coherence) is the OECD as a distinctive organisation that is highly adaptive, expansionist, and self-serving. At the same time, however, it continues to cling to much from its past, as it attempts to forge a new identity and future. As such, legitimacy and organisational survival are essential factors of concern.

---

[5] International Institute for Educational Planning, an arm of UNESCO created in 1963.

Table 3.4 Key literature featured.

Theme-title	Author-date
OEEC, OECD educational activities (1950s-1990s)	
1. *30 Years of Educational Collaboration in the OECD*	Eide (1990)
2. *Education 1960-1990: The OECD Perspective*	Papadopoulos (1994)
3. *Educational Multilateralism and World (Dis)Order*	Mundy (1998, 2007)
4. *Global Governance, Educational Change*	
5. *The OECD and Cold War Culture: Thinking Historically about PISA*	Tröhler (2013)
6. *Learning Productivity: The European Productivity Agency – An Educational Enterprise*	Bürgi (2019)
7. *The OECD, American Power and the Rise of the "Economics of Education" in the 1960s*	Elfert (2019)
8. *The Birth of the OECD's Education Policy Area*	Centeno (2019)
OECD educational activities and changing approach (1990s-2010)	
1. *International Comparative Studies in Education: Descriptions of Selected Large-Scale Assessments and Case Studies.*	National Research Council (1995)
2. *The OECD, Globalisation and Education Policy*	Henry et al. (2001)
3. *How to Become an Influential Actor – The 'Comparative Turn' in OECD Education Policy*	Martens (2007)
4. *The OECD and Global Shifts in Education Policy*	Rizvi and Lingard (2009)
5. *The OECD and the Expansion of PISA: New Global Modes of Governance in Education*	Sellar and Lingard (2014)
OECD post-2011 governance and global expansion, PISA-D (2011-2030)	
1. *Expert Moves: International Comparative Testing and the Rise of Expertocracy*	Grek (2013)
2. *IIEP Staff Interview with Michael Ward on PISA-D*	UNESCO IIEP (2016)
3. *Camilla Addey Interview with Andreas Schleicher on PISA and PISA for Development*	
4. *Golden Relics & Historical Standards: How the OECD is Expanding Global Education Governance through PISA for Development*	Addey (2016a, 2017)
5. *PISA for Development: How the OECD and World Bank Shaped Education Governance Post-2015*	Auld, Morris and Rappleye (2019)
6. *The Organisation for Economic Co-operation and Development's International Early Learning Study: What's Going on.*	Moss and Urban (2019)

Data interpretation: the hermeneutic circle–spiral

As already discussed in the explanation of the research design, at the core of this methodology lies a critical hermeneutic interpretation. That means that the interpretive process is governed by a hermeneutic circle which leads towards deeper and richer understanding (Dilthey 1976; Gadamer 1976; Bentz and Shapiro 1998). For Schwartz-Shea and Yanow (2012), it expresses the idea that there is no fixed start/end point for the inquiry but one continual interpretive path which can be better conceived of as a spiral rather than as a circle. In this respect, the hermeneutic circle–spiral characterises the interpretive/learning process in general. It also describes the intimate relations between the part and the whole, highlighting how the meaning of an individual text or act relates to the whole, while the overall coherence of argumentation is inseparable from its constituent parts. Kelly (2002), in her inquiry into the decision-making process in cases of child protection within organisations in the UK, summarises the hermeneutic process as:

> …immersing oneself in the data, allowing themes or subthemes to emerge, considering the themes in relation to the overall situation and 'standing' of the documents and a constant reflectivity on the interpretation of the text in relation to other forms of data. (191–192)

In this study, the critical hermeneutic interpretive approach has two major uses. The first is to refer to the overall interpretive logic of this study which is explained in the description of the research design (see Figure 3.1) and is closely linked to the idea of intertextuality – reading across different forms of data. The second is to specifically describe the processes of interpreting OECD documents to elicit meaning and gain understanding. The application is twofold because in interpretive research, the process of interpretation itself is the instrument. Understanding and interpreting (through intertextuality and reflexivity) are intrinsic to the research process.

To apply the critical hermeneutic method to analyse organisational documents, Forster (1994) suggested seven stages (reproduced in Figure 3.2). The 1st stage involves understanding the meanings of individual texts; at this stage, the focus is to carefully read each document and to draw out the 'taken-for-granted' assumptions and viewpoints of their authors (in the context of this study, the OECD and to a lesser extent the PISA-D countries). This requires attention to the 'units of relevant meaning' (such as those shown in Table 3.1) which appear in the texts. The 2nd stage is to identify how these different themes are related to each other and to see if there is a central theme emerging which

provides a higher-order level of understanding of sub-themes (e.g. the OECD's highly evolved capacity, discussed in Chapter 4; concerted efforts to legitimate PISA-D in Chapter 5; and self-serving needs embedded in PISA-D outcomes in Chapter 6). Based on the extraction, the 3rd stage is to group texts that have an inner unity and commonality of meaning. It then becomes possible to cluster documents according to their own inner cohesion and logic (e.g. those clustered in the 4th column in Table 3.1; the same documents containing different texts are not repeated in the table).

Once these 'clusters of meaning' have been obtained, the 4th stage is to relate them to the research question(s). Here, Forster stresses that the true meaning of a text can only be elicited with reference to other texts and other forms of data (in this study, the interviews and the literature), and 'the meaning of individual textual segments can be very different from those that would emerge if the entire document were used as the basis for analysis, or if all the data at hand were employed' (ibid., 152). In response to this, the 5th stage is then to employ reliability and validity checks to attain the intellectual rigour while interrogating these texts. However, this evaluative stage is one where disagreement remains. Schwartz-Shea and Yanow (2012) argue that both 'reliability' and 'validity' are commonly accepted positivist standards which do not fit interpretive methodological concerns. They suggest focusing instead on 'the adequacy of explanation and analysis – the explanatory coherence of the argument' (ibid., 108). Nevertheless, all the authors point to (i) the consistency of evidence from different sources, and (ii) engagement with conflicting or contradictory interpretations. Thus, according to Forster (1994), if the 'reliability' and 'validity' checks are successful, the 6th stage is to recontextualise these data within the broader organisational contexts and processes and with reference to other types of analyses. The final stage is to present case materials from the documents that have been examined.

I have adapted and extended Forster's (1994) seven-stage hermeneutic process to analyse the OECD documents, taking into account Schwartz-Shea and Yanow's (2012) points about intertextuality, logic of argumentation and explanatory coherence. These changes are implied at stages 4 and 5 in Figure 3.3. The additional 7th stage of *enacting reflexivity* considers self-conscious reflections on my sense-making during the analysis, including the themes identified and the explanations and argumentation involved. Reflexivity also involves going backwards in time to contemplate initial theoretical expectations as my understanding evolves and deepens, and forward as I ponder

emerging inconsistencies or even contradictions (e.g. between OECD's proclamation and findings emerging from the national reports – there has been little empirical substantiation for the idea that PISA-D results provide national policy makers with data and evidence that can be used to improve the quality of education; Chapter 6). Moreover, Schwartz-Shea and Yanow (2012) suggest that interpretive researchers not only reflect on their own meaning-making, but also on what they are *not* hearing, and on the silences in their interviews and close readings of different materials. From a reflexive sociological point of view (e.g. Bourdieu and Wacquant 1992, 40), this is related to exploration of 'the unthought categories of thought which delimit the thinkable and predetermine thought'. Accordingly, while sequentially it occupies a separate stage, reflexive checks run through the hermeneutic process, including the concluding chapter where I reflect on the overall interpretive journey (Chapter 7). Chapters 5 and 6 present the case materials in greater detail.

Figure 3.2 Forster's (1994) hermeneutic process.

(Research design)

| 1 Understanding the meanings of individual texts |
| 2 Identifying (sub-) themes |
| 3 Identifying thematic clusters |
| 4 Triangulating documentary data |
| 5 Employing reliability and validity checks |
| 6 (Re)contextualizing documentary data |
| 7 Using representative case material |
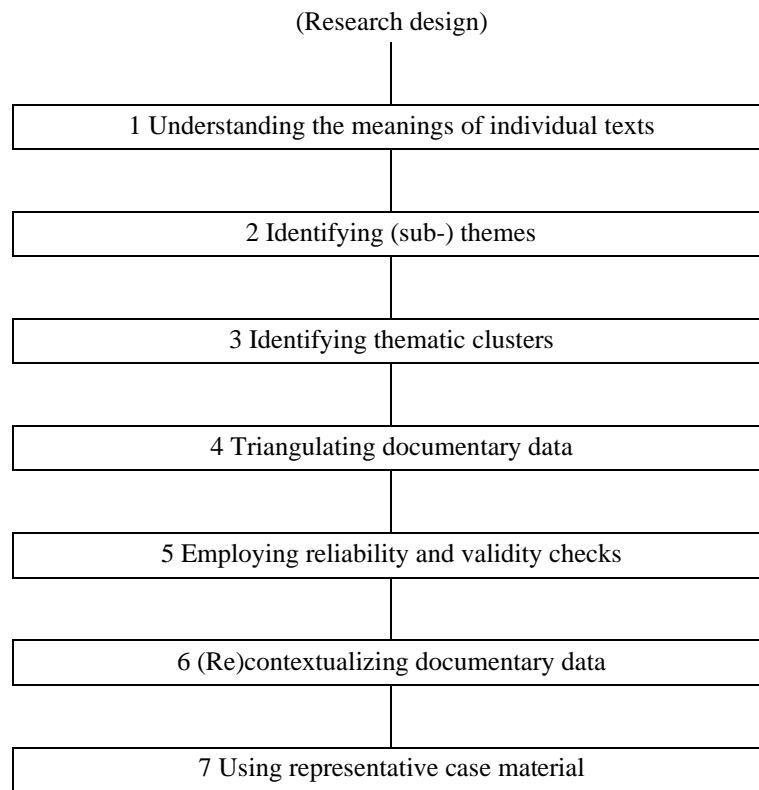
Figure 3.3 The hermeneutic process – *adapted*.

(Interpretive research design)

| 1 Understanding the meanings of individual texts |
| 2 Eliciting (sub-) themes |
| 3 Identifying thematic clusters |
| 4 Reading intertextually across documents, interviews and literature |
| 5 Checking coherence and logic of argumentation |
| 6 (Re)contextualizing documentary data |
| 7 Enacting reflexivity |
| 8 Using representative case material |
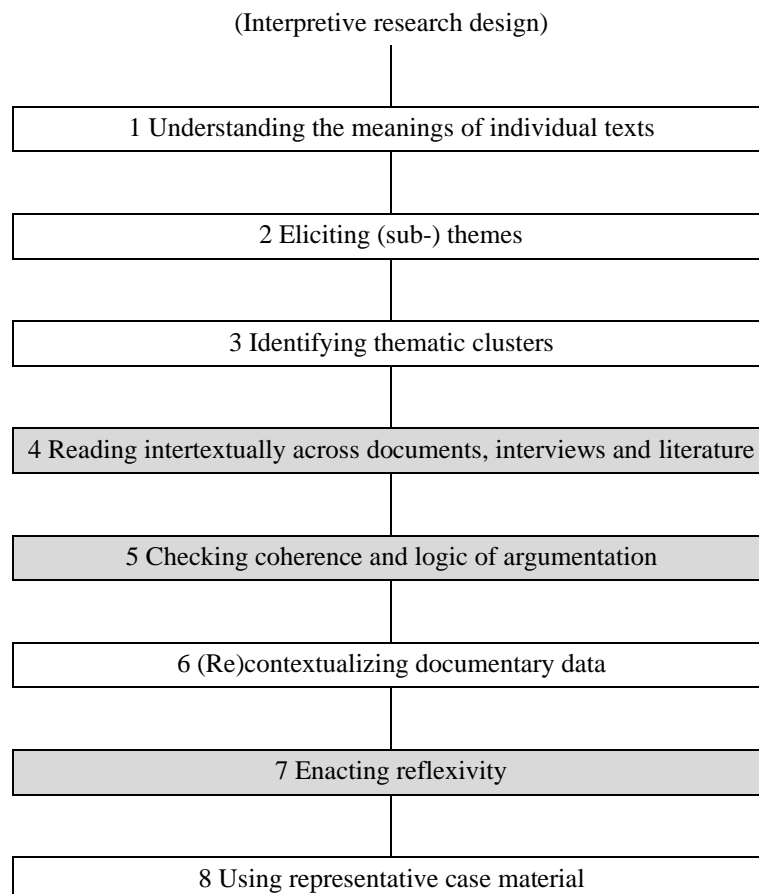
Ethical considerations and trustworthiness

Ethical issues could arise at various stages of social research revolving around relations between the researcher and her research participants (Bryman 2012). Diener and Grandall (1978) categorised these issues under four aspects: (i) whether there is harm to participants; (ii) whether there is a lack of informed consent; (iii) whether there is an invasion of privacy; and (iv) whether deception is involved. However, these are generic ethical principles for social research, and in this study, they need to be considered within the rhythm of interpretive research.

As discussed in the preceding sections, interpretive researchers generally have limited control over the research settings and processes. This is because the interpretive researcher immerses herself in the field realities and co-generates the data through interacting with her research participants (or documents). In this study, the lack of control was clearly observable as I approached (and in many cases was turned down) and interviewed organisational and political elites and experts. Almost all interviews were carried out in ways that accommodated the preferences of my participants, including time, location, and whether over the telephone or face-to-face. They were informed in advance of the kinds of questions they would be asked and assured that they could choose to answer certain questions or end the interview whenever they wished. The risk to my participants was also minimal considering their elite status and public roles (for instance, as directors or senior policy analysts at the OECD or within countries). The answers they gave represented their general positions. Moreover, it has been argued that 'elite groups are more difficult to penetrate than other groups…[and] are better equipped to protect themselves' (Smith 2006, 643).

Unlike other types of interpretive research methods, such as participant observation, in which researchers struggle with how best to safeguard participants' identities and mask their sites (see Buzzanell 2018), this study used mainly the OECD documents which are already available in the public domain. In fact, the OECD itself has recognised the great research potential of publicly available data for yielding improved insights into society and political behaviour. In recent years, it launched the Open Government Data[6] (OGD) initiatives to promote transparency, accountability and value creation by making

---

[6] http://www.oecd.org/gov/digital-government/open-government-data.htm

government data available to all. In the field of education, PISA itself is a mechanism for data generation and a powerful policy tool, which the OECD claims has helped to change the balance of power in education by making public policy more transparent and more efficient (see e.g. Schleicher 2013b). Therefore, it would be at odds with the OECD's own belief and initiatives if the documents it has listed on the PISA-D webpage cannot be used for interpretive research.

In terms of power, scholars have openly questioned and called for the OECD's commitment to transparency and public accountability in its education decision-making process; thus, the *Open Letter to Andreas Schleicher* (2014, 873) states, 'an organisation like OECD, as any organisation that deeply affects the life of our communities, should be open to democratic accountability by members of most communities'. This implies openness on how the OECD and its private contractors assess the education of 15-year-old students (along with 14–16-year-old out-of-school youth) and with what outcomes should be subject to public scrutiny. In applying the critical hermeneutic approach, this study takes on the responsibility to uncover what has been claimed as scientific rationality and data objectivity. I thus agree with Mølstad and Pettersson (2019) that nowadays quantifications and comparisons represented as numbers have become a more or less hegemonic way of telling the 'truth' about education, and we need to refrain from a single reading.

Lincoln and Guba (1985) introduced the notion of 'trustworthiness' as a tool of assessment for judging the overall quality of a study. Trustworthiness differs from the positive standards of 'validity' and 'reliability' as it emphasises the humanistic aspects of interpretive research (Yanow and Schwartz-Shea 2015). As Riessman (2002, 258) puts it, 'validity' and 'reliability' assume an objective reality; trustworthiness 'moves the process into the social world'. According to Schwartz-Shea (2006), trustworthiness concerns the many steps that a researcher takes during the course of an inquiry to ensure that her efforts are self-consciously deliberate, transparent, and ethical. These steps entail evaluative criteria, such as 'thick description', reflexivity, and consistency with interpretive presuppositions. While 'thick description' is more often used in ethnography (Geertz 1973), it elicits the central theme of the hermeneutic–phenomenological approach: embedding meaning in context. Forster (1994, 164) also emphasises the need for 'thickly crafted', situated documentary reading:

> The usefulness of these data depends greatly on the care and skill employed in their handling, since many of these cannot be taken at face-value. They can never be analysed in isolation and can only be fully understood in the context of a holistic view of the organisation and of other types of analyses.

This study provides a 'thick description' of PISA-D, firstly by situating it within an analysis of the OECD's long-term evolution and development, and secondly by reading across different data sources (including documents, interviews and literature) to pursue the intertextuality of analysis. Reflexivity is another hallmark of interpretive research concerning self-conscious reflections on sense-making processes and outcomes. In this study, I have incorporated reflexive checks into the hermeneutic process of interpreting OECD documents (see Figure 3.3) and in the last chapter where I revisit the major outcomes and reflect on my positionality. Schwartz-Shea and Yanow (2012, 100) appreciate such efforts, noting that '[e]ngaging in these practices and making one's engagement explicit and as transparent as possible in the research [design as well as research] manuscript is understood within the interpretive epistemic community as contributing to the quality of interpretive research'. Clearly, these reflexive practices will perturb positivist researchers who presuppose a single absolute account of social reality (pursued over a programmatic research strategy). These reflexive checks demonstrate their interpretive consistency.

Chapter 4. PISA-D Origins and OECD Evolution

The new men of Empire are the ones who believe in fresh starts, new chapters, new pages; I struggle on with the old story, hoping that before it is finished it will reveal to me why it was that I thought it worth the trouble.

John M. Coetzee, *Waiting for the Barbarians*

They are political actors in their own right, having their own particular resources for shaping political action, and both shaping and being shaped by others.

Michael Barnett and Martha Finnemore, *Rules for the World: International Organisations*

Introduction

In this chapter, I trace the origins of PISA-D and interpret the processes by which they have shaped its contemporary form. The argument is that PISA-D directly addresses the core elements of the OECD's strategic vision for 2011 and beyond, and needs to be understood as part of a long-term pattern of adaptation, survival and expansion as the OECD both responds to changing global realities and shapes the niche in which it has evolved. Based on this understanding, I have developed a periodisation of the OECD's changing agendas and approaches to education, in which these dimensions can be better illuminated (they are summarised in Table 4.1). The analysis is informed by the constructivist approach which perceives IOs as autonomous sources of power that both constitute and construct the social world (Barnett and Finnemore 1999, 2004). The adoption of a timescale – from the 1950s to 2030 – provides further insights into the OECD's continuous endeavour to reinvent itself and increase its power and relevance, which, in the post-2015 context, has both manifested and culminated in the OECD's strategic work in developing PISA-D and engaging with the SDGs.

Overall, as a road map, I have identified three distinctive time periods. The first is the period from the 1950s to the 1990s during which the OECD was involved in Cold War politics, and its education agenda (which had formed the ideological basis of PISA and PISA-D) mainly served geopolitical and economic purposes. The second period was from the 1990s to 2010, and was characterised by processes of globalisation and the move towards a global knowledge economy. Within this context, the OECD shifted its approach from managing countries individually to developing international comparisons of

educational outcomes (e.g. the birth of PISA) and setting an overarching framework for monitoring education systems (i.e. DeSeCo). The second period shows that the OECD assumed an increasingly prominent role in global educational governance. The third period starts from 2011, when the OECD set out a new strategy for development, and continues into the year 2030. I show that during this period the OECD assumed a dominant position whereby it can continuously shape the future of education and development. PISA-D, specifically, has been essential for realising the OECD's post-2011 agenda by bringing low- and middle-income countries within its realm of control (through a 'larger' PISA programme) and accelerating the global shift from learning for all to 'assessment for all'.

It is important to recognise that for the periodisation used in this chapter, there is never a clear-cut boundary between the identified periods wherein the OECD's agendas and approaches to education suddenly change. Transition is a slow process in which certain aspects are increasingly emphasised and others, having served their initial purpose, gradually become less prominent, or take on an intrinsic value (such as the 'economics of education' approach). The use of periodisation here is more of a heuristic device – a 'tool of thought, valid in so far as it is illuminating, and dependent for its validity on interpretation' (Carr 1964, 60). Following Carr, I develop the periodisation in this chapter to both illuminate the OECD's highly evolved capacity and expansionist tendency in changing geopolitical and economic contexts and to provide a critical understanding of PISA-D against this background. In doing so, this chapter will also provide a solid foundation for the analysis that follows of the PISA-D legitimation strategies and outcomes. I provide a detailed interpretation below of each of these periods.

Table 4.1 Summary of the main features of the OECD's changing agendas and approaches to education.

Context	Cold War Politics	Neo-liberal Globalisation	Sustainable Development Goals
	1950s – 1990s	1990s – 2010	2011 – 2030
Clients	Industrialised countries (18 European countries + US and Canada)	Industrialised countries and emerging countries (e.g. the BRICs + South Africa)	Industrialised countries, emerging countries (e.g. the BRICs + South Africa) and developing countries in Africa, Asia, Latin America + the Caribbean
Agenda	An inferred role for education (science and technology)	Economic-oriented (education for economic growth)	More comprehensive view of education and development (education for sustainable development)
Ideology	Productivity boost and economic recovery (capitalism)	Human/Knowledge capital formation and global competitiveness (neoliberalism)	Inclusive growth and systematic thinking (banal imperialism)
Role / Influence	Catalyst/facilitating (technical-rational expertise)	Sculptor/shaping (policy pathfinder & standard setter)	Arbiter/decisive (shaping global architecture on education and development)
Approach	Sharing information and acknowledging differences	Scientific approach and 'comparative turn'	Comprehensive approach and 'humanitarian role'
Example	Country reviews and report on basic educational statistics of individual countries	Cross-national comparisons (INES, PISA) and an overall conceptual framework for managing education systems (DeSeCo)	Global learning metric (PISA-D, Learning Framework 2030) and PISA as part of a 'configuration' that provides multidisciplinary expertise and a governance toolkit which can be tailored to all countries

From the 1950s to the 1990s: Cold War politics and education for economic recovery and growth

The OECD's predecessor, the OEEC, was established in 1948 to administer the U.S.-funded Marshall Plan for European economic recovery after the Second World War (1939–1945). Eighteen Member countries[7], primarily based in western Europe, initially joined the OEEC. Its reconstructive work was significant in that it had laid the ideological and institutional foundation for many of the OECD's earlier activities in the sphere of education. In particular, the European Productivity Agency (EPA), which was set up in 1953 as a semi-autonomous operational body under the OEEC, aimed to maintain Europe within the ideology of the capitalist West and to boost the productivity of its economies. According to Danish historian Bent Boel's (2003, 9) historical account,

> [t]he goal of the productivity program in Europe is nothing less than the reshaping of the European economies from a static pattern characterized by restrictionism into a pattern of dynamic and expanding free and competitive enterprise. This alone will make possible continually rising living standards and ever-increasing consumption of more and better things by more people. Behind this objective, of course, and the ultimate justification, is the compelling need to keep the countries of Europe willing and effective partners in the free world. (The U.S. FOA [Foreign Operational Administration] Productivity Programme Memo, 1954)

Within this 'reshaping' process, the educational dimensions of the EPA played a key role in changing, first and foremost, European attitudes towards productivity. These included, for example, embracing technological change and ideas associated with exploitation and commodification. The major strategy in this regard was to 'educate' change agents, such as managers, union leaders, officials, and workers, who were best placed to influence the attitudes of others in large numbers (Bürgi 2019). To this end, the EPA/OEEC also established a close collaboration with the Ford Foundation (established in 1936) to develop management education in western Europe and to foster the spread of American managerial skills and techniques (Gemelli [1996], cited in Elfert 2019).

The productivity drive also led to the maximisation of the relationship between technological-scientific development and economic growth. An example of this is the creation of a committee for technical and scientific personnel (i.e. the Governing Committee for Scientific and Technical Personnel) within the OEEC in 1958. Alexander

---

[7] Austria, Belgium, Denmark, France, Greece, Ireland, Iceland, Italy, Luxembourg, Netherlands, Norway, Portugal, Sweden, Switzerland, Turkey, United Kingdom, Western Germany, and the territory of Trieste (joined in 1949).

King, who played a major scientific role in the British war effort during the Second World War and was appointed as the deputy director of EPA in 1957, was in charge of this area of work. According to Bürgi (2019, 22), King, with a background in chemistry, 'was a social engineer par excellence and very much engaged in transferring war planning techniques (such as operations research) to the social sciences'. He also 'belonged to a circle who saw scientists as the key political actors in the future society, and he wanted the OEEC to promote science and technology' (Eide 1990, 8). The ground was thus prepared along these lines for the particular features of the OECD's work in education in its early decades.

By the end of the 1950s, with the mandate for reconstructing Europe fulfilled, a resolution was adopted by its leaders to adapt the OEEC to perform a wider role. The OECD thus came into being in 1961 as a *transatlantic* platform that worked to shape policies and promote economic growth in industrialised western countries (which also included the U.S. and Canada). The extended mandate also brought in new fields of collaboration. For example, with the emerging movement for decolonisation in the late 1950s, emphasis was given to coordination of Member countries' policies towards the developing world through the Development Assistance Committee (DAC, founded in 1960 under the Development Co-operation Directorate [DCD]; Eide 1990; Papadopoulos 1994; Mundy 2007). However, Hongler (2017) observes that the DAC worked as a place where the OECD countries positioned the West as a competent centre of rationality versus an emotional and needy Global South. She conceives this kind of positioning as a specific way of the West dealt with the threat of decolonisation and for legitimating colonial rule. I shall demonstrate later in this chapter that the DAC/DCD played a critical role in supporting the extension of PISA into such nations, primarily through approaching development partners (such as donor countries and leading agencies) and initiating dialogues with them about PISA-D. With regard to the transformation of the OEEC to the OECD, Ydesen and Grek (2020, 4) consider that it was an outcome of an organisational 'struggle for survival' in changing historical contexts.

The OECD did not have a mandate for education which can be identified in the OECD Convention. According to Papadopoulos (1994, 11), '[t]he nearest it comes to getting such a reference is in Article 2*(b)*, on policies designed to promote the development of Member countries' resources in science and technology, encourage research and promote vocational training'. As mentioned earlier, the OECD became involved primarily because

of Cold War politics and the influence of American power. Specifically, there were two major events or processes that contributed to the formation of the OECD's education agenda. The first was the Sputnik shock and the Space Race, which refer to the U.S. response to the Soviet Union's launching of the first man-made satellite, 'Sputnik-1', into space in 1957. According to Tröhler (2013, 144), before Sputnik, 'the US authorities were completely certain that the United States had taken the global lead in education, science and technology'. The Sputnik shock led to two major reactions: the founding of the National Aeronautics and Space Administration (NASA) and the reshaping of the education system. As former U.S. President Eisenhower (1958, 101) declared,

> [t]he Nation requires the fullest development of the mental resources and technical skills of its young men and women. The present emergency demands that additional and more adequate education opportunities be made available. The defense of this Nation depends upon the mastery of modern techniques developed from complex scientific principles … This requires programs that will […] correct as rapidly as possible the existing imbalances in our education programs which have led to an insufficient proportion of our population educated in science, mathematics, and modern foreign languages and trained in technology.

The second related development underpinning the OECD's involvement in education was the emerging view of education as an important source of economic growth (therefore, a predictable investment in human capital), which was known in the U.S. as the economics of education. A culmination of this view was the 1960 Presidential address to the American Economic Association given by the economist (and later Nobel prize winner) Theodor Schultz. He (1961) declared that traditional economic factors, such as land, man-hours and physical reproducible capital, were not sufficient to explain the annual growth of the national product, and then showed that the educational level of a country could be an important factor in this aspect. Accordingly, the idea was soon adopted by the OECD and followed up through the creation of the Study Group on the Economics of Education (Eide 1990; Holden and Biddle 2017). Papadopoulos (1994, 33) comments that the Study Group 'marks the beginning of the consecration of the economics of education not merely as a new discipline, but as an area of direct policy relevance and concern'. Similarly, Elfert (2019, 40) observes how these new studies in a climate of economic growth, along with the worsening of the Cold War and the Sputnik shock, 'triggered a veritable frenzy of educational research and planning activities that opened up a new field for the OECD' (see below).

As a result of these major factors and the adaptive capacity of the OECD, the 1960s saw it becoming involved in activities of 'educational collaboration [and] exchange of

experience and ideas' (Eide 1990, 1). The focal point was on the collection of statistical data on technical and scientific personnel in Member countries, along with country reviews of the educational systems supporting these fields. This area of work was carried out under the Committee for Scientific and Technical Personnel (CSTP). Papadopoulos (1994, 37) explains that the CSTP 'was quick to cash in on new opportunities opened up by the economic growth objective', which was in fact 'an inspired, intentional move designed to secure a lasting place for education within the new Organisation'. In addition, the OECD expanded its activities also in educational planning, especially the Mediterranean Regional Project (MRP) on manpower forecasting. According to Eide (1990), the MRP offered opportunities for the OECD not only to help raise the profile of national planning bodies within its Members countries but also to exert a degree of influence on national educational policies. The following quote from Eide is an illustration of how the OECD exercised its technical-rational authority:

> Educational policy became interesting for the OECD because of its assumed importance for economic growth and development. The planning project in the Mediterranean was clearly based on this. The national groups which were created, had their prime contacts with national economic authorities, and especially newly established planning ministries. Such ministries regularly were in conflict with the ministries of education, which they saw as outdated and uninterested in economic issues. For the OECD, it was easy to get into those countries, because the organization seemed to legitimate the right of economic authorities to intervene in educational policy. Probably, the OECD secretariat also used available means for technical cooperation as an encouragement to the personal alliances formed in the member countries. The latter thus got the means for implementing their own national pet projects. (1990, 15)

The CSTP was later transformed into an Education Committee in 1970, with a wider mandate to address education policy issues. With a grant from the Ford Foundation, the Centre for Educational Research and Innovation (CERI)[8] was established in 1968 to focus on educational research. Centeno (2019) suggests that CERI was created partly in response to the perceived need for some qualitative research within the OECD to complement the Education Committee's more quantitative and descriptive work; moreover, this endeavour also provided a platform for the OECD Secretariat to secure the continuation of the OECD's educational activities and to put forward a concrete educational programme. These dimensions of continuity and survival can also be observed from Eide's (1990, 31) account that 'the two bodies [i.e. the Education Committee and CERI] would have been amalgamated [a] long time ago, if it had not been

---

[8] Andreas Schleicher joined the Centre in 1994 as a project manager.

for the fact that their existence secures more money for educational activities in the OECD'.

Due to the oil crisis in the mid-1970s, many OECD Member countries faced high unemployment rates, especially among young people. Consequently, in the 1980s, the OECD's educational activities were mostly concerned with employment. For example, "[b]oth the Education Committee and CERI developed extensive projects concerning the transition from education to work, the interplay between education and labour market policy, the "youth guarantee" of the Nordic countries, etc.' (Eide 1990, 43). Overall, the basic principle of the OECD's approach to education in its earlier decades was the recognition of education policies as part of national policies reflecting the specific realities and demands of individual countries. While the OECD was also able to exert some influence on education policies of Member countries, according to Eide (1990, 16), countries were often able to 'effectively block planning initiatives from the OECD, and to prevent any real impact on educational policies'. As such, the OECD policy ideas tended to focus on emerging issues on the education horizon, and the framing of these issues involved a dialectical process with exchange of ideas within individual countries. The outcomes were then fed back to the countries to be followed up based on national circumstances as well as interest and capacity (Papadopoulos 1994).

Therefore, the OECD's influence on Member countries must be understood as a *catalytic* role that related to identifying major policy issues and facilitating policy making, programme planning and implementation; this remained the key feature of its work in education from the 1960s through to the 1990s (Papadopoulos 1994). This specific period, as already noted, was the time when a bipolar world order was constructed, and international organisations served the geopolitical interests of the two hegemonic powers (Mundy 1998, 2007). Consistent with the conventional realist perspective (described in Chapter 2), the OECD operated under the U.S. leadership. Therefore, in addition to supporting economic cooperation and development, the OECD also played an important role in fighting against communism (Eide 1990; Tröhler 2011, 2013). The fact that the member countries of the OECD was limited in number and included mainly market-based economies enabled its tasks to be managed more easily and at a level of sophistication that would not have been possible if membership was wider (Papadopoulos 2011). This has implications for the ensuing change in the organisation's approach to education during the post-Cold War period of neoliberal globalisation.

From the 1990s to 2010: neoliberal globalisation and development of international comparisons

The OECD had to change its approach to education as its membership expanded and diversified, along with investment in new partnerships with non-member economies. According to the *Acts of the Organisation* (Vol. I), in 1990, the Executive Committee consisted of Permanent Representatives of Australia, Austria, Belgium, Canada, Denmark, France, Germany, Italy, Japan, Portugal, Sweden, Turkey, the UK and the U.S. Leimgruber and Schmelzer (2017, 8) suggest that the aim of the organisation during this transitional period was to reposition itself from a (western) European organisation into a 'triadic [Euro-Atlantic-Pacific] capitalist think tank'. As such, the 1990s saw the OECD find its role *vis*-à-*vis* the European community, 'primarily as a link between industrialized countries inside and outside Europe, [with] less political control, and less faith in specific economic/political ideologies' (Eide 1990, 52). However, Rizvi and Lingard (2006, 2009) observe the tensions and infighting between the social democratic (endorsed by the European members) and the neoliberal perspectives underpinning the organisation's policy frameworks and conclude that in the context of globalisation and with the global hegemony of the U.S., the neoliberal paradigm dominates.

Within education, this neoliberal paradigm has expressed itself through the reframing of education as central to national economic competitiveness, driven by human capital discourses and linked to an emerging global knowledge economy (OECD 1996b, 2004; Rizvi and Lingard 2006; Grek 2009; Kallo 2009). The increased importance of education was also recognised and reflected in a new institutional structure at the OECD. In 2002, a separate Directorate for Education was established to elevate the profile of education work within OECD countries. At the technical level, this paradigm has required the collection of statistical data to say more about outputs or outcomes of education than about its inputs and contexts (also see Ball 2003). 'The OECD paradigm is very much about making learning outcomes visible, [and] learning metrics around them' (Interview OECD #2, 2018).

In a retrospective reading, these dimensions are encapsulated in the vision statement of the OECD Ministerial Level Meeting (in 1990), in which, as the organisation ushered

in a new decade, it called for a reformulation of educational purposes and governance in line with the requirements of the global economy. As it stated,
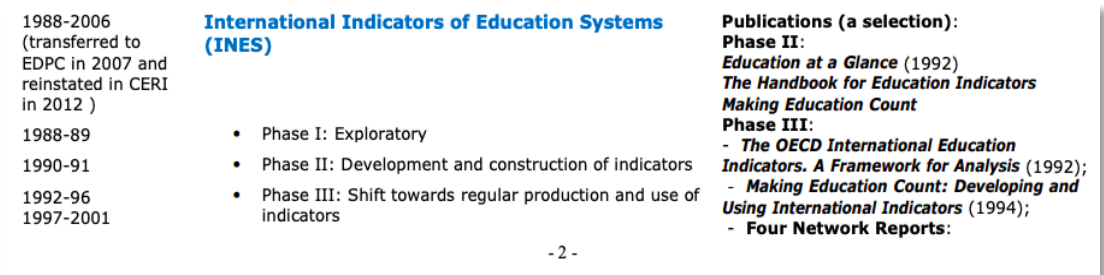
> [i]n recent years, many OECD countries have sought to raise the quality of schooling as a policy priority. Criticisms and concerns have added urgency to the task: fears that too many youngsters have been leaving education *ill-equipped* for the modern world, that *standards*, however they compare with the past, *need to be higher still*, that educational deficiencies might lead to slippage in *worldwide economic competitiveness*.

> A poor educational foundation becomes exacerbated on *entry to the labour market*, and while the previously very high levels of youth unemployment are tending to decline so too are the numbers of low and unskilled jobs – *knowledge and skills are today at a premium…* These reinforcing patterns call for *fresh solutions* – in mainstream provision through *targeted measures* and through the creation of new opportunities in cooperation with other economic and social partners. (OECD 1990, 4–6, my emphasis)

Against this background, Martens (2007) identifies a significant shift in the OECD's approach to addressing educational issues which is related to a deliberate and increasing reliance on *ranking and rating* countries with comparative indicators. She argues that this was operated through a 'comparative turn' – 'whereas in the past it has focused on each state individually acknowledging differences and idiosyncrasies, it now decisively compares states with each other and against standardised criteria' (ibid., 2007, 40). The 'comparative turn' was greatly facilitated by the OECD International Indicators of Educational Systems (INES, 1988–2006) project managed by the CERI. It was a deliberate effort to develop a set of comparative education indicators for cross-national comparisons and provide a platform for their use in national policymaking and managing education systems (National Research Council 1995).

The initial work of the INES project was carried out by a consortium of countries comprising a Technical Group and four Networks: Network A – Educational Outcomes (chaired by the US); Network B –Education and Labour Market Destinations (chaired by Sweden); Network C – Features of Schools (chaired by the Netherlands); and Network D –Attitudes and Expectations (chaired by the UK). These Networks were coordinated through a structured framework, which chronologically involved three phases of implementation (a brief timeline is reproduced in Figure 4.1).

Figure 4.1 Timeline of the INES project and major publications.

1988-2006 (transferred to EDPC in 2007 and reinstated in CERI in 2012)	**International Indicators of Education Systems (INES)**	**Publications (a selection):** **Phase II:** *Education at a Glance* (1992) *The Handbook for Education Indicators* *Making Education Count*
1988-89	• Phase I: Exploratory	**Phase III:** - *The OECD International Education Indicators. A Framework for Analysis* (1992);
1990-91	• Phase II: Development and construction of indicators	- *Making Education Count: Developing and Using International Indicators* (1994);
1992-96 1997-2001	• Phase III: Shift towards regular production and use of indicators	- **Four Network Reports:**
	- 2 -	

Source: OECD (2012c), *Historical Summary of CERI's Main Activities*

Accordingly, Phase I of the project explored the feasibility of developing and reporting a key set of indicators that had been in use by Member countries. These included student flows, student outcomes, functioning of schools, costs and resources, and attitudes and expectations. Each Network had a lead country coordinating the contributions of Members, supporting the theoretical and technical work required, and producing the Network report. For example, Network A, which was led by the U.S. Department of Education, National Centre for Education Statistics, was responsible for the development and preparation of educational outcome indicators.

In Phase II, the Networks selected some initial indicators (including the context within which education systems operate, costs, resources and school processes, and outcomes of education), established data sources, and calculated some experiential indicators. The results, analyses and requirements for follow-up work were presented and published in *Education at a Glance* (the first edition came out in 1992). This process was guided by a small group of expert advisors who played a key role in bringing the exploratory phase to a satisfactory conclusion (Papadopoulos 1994). In the third and final phase, the aim was to establish an institutional framework to facilitate the regular production of a set of INEs (which became part of the regular responsibilities of the OECD's Indicators and Analysis Division, Directorate for Education and Skills). The third phase also included continuation of ongoing conceptual work and regular publication of the calculated indicator set as *Education at a Glance*.

When the INES project was completed, the OECD and its Member countries succeeded in 'setting up an international database containing up-to-date and relevant education statistics and indicators; a fully computerized on-line network for collecting and disseminating the data; and cost-effective procedures for processing the data'

(National Research Council 1995, 39). Intertextually, these outcomes support Martens' (2007, 46) view that with this new approach to education statistics, 'the nature of OECD indicators changed significantly in quantitative and qualitative terms: they became more regular, more reliable and more meaningful sources of data'. Compared to the period before the 1990s when this development had proved to be premature (Papadopoulos 1994), the renewed interest and resulting institutional structure of INES encouraged systematic and regular publication of educational statistics (Martens 2007).

However, as Papadopoulos (1994, 191) anticipated decades ago, while *Education at a Glance* provided a handy source of reference for cross-national comparisons of educational performance, it was 'only the beginning of […] a long and arduous process of developing a complete framework of international education indicators'. Since then, this endeavour has been continued and sustained through the OECD programme *Definition and Selection of Competencies: Theoretical and Conceptual Foundations* (DeSeCo, 1998–2002) to formulate a coherent strategy on defining, selecting and measuring competencies and skills (OECD 2001). In other words, the programme sought to enhance the INES project (especially Network A) by defining its underpinning principles and to provide guidelines for future development on the evaluation of education systems and of learning outcomes, in particular. The following are the major questions that DeSeCo aimed to address:

> What competencies – apart from reading, writing, and computing – are necessary for individuals to lead an overall successful life and for society to face the challenges of the present and the future? With the myriad of conceivable individual and social differences in and across countries, is there any ground for defining *a general (or universally applicable)* set of key competencies? And if so, based on what normative, definitional and conceptual criteria can a *limited* set of key competencies be identified? Such questions were at the heart of OECD's interdisciplinary and policy-oriented research program DeSeCo […] launched at the end of 1997 as part of OECD's INES Education Indicators Program. (OECD 2003b, 2, my emphasis)

The agenda of DeSeCo responded to the changing governance approach enacted within the OECD from the 1990s. As touched upon earlier, the neoliberal paradigm placed a stronger focus on outcomes and performance in relation to efficiency, effectiveness, and quality of services, and accelerated a shift in the central management bodies towards setting the *overall* framework rather than micromanaging (OECD 1995; Ball 2003). The changing dynamics were well reflected in the CERI's educational programmes, which the former Head, Jarl Bengtsson (2008, 2), summarises: 'From the 1990s to the present day, many of the same preoccupations continued but there were

important new priorities. These included the focus on *learning*; *knowledge* – the knowledge economy and knowledge management; and the *international* dimension' (original emphasis). In this respect, DeSeCo was critical for advancing the development of a common, overarching framework for monitoring education systems and managing knowledge among OECD countries and beyond. Read intertextually, Grek (2013, 703) notes that it 'moved policy directions and thus policy-makers and nations towards a common, coherent international discourse on competence and skill development'. Consequently, the framework has also enabled educational experts to *disengage* from 'the myriad of conceivable individual and social differences in and across countries' (OECD 2003b, 2) leading to the establishment of a form of central steering.

In terms of measurement and cross-national comparisons, the most direct application of the DeSeCo conceptual framework was the development of PISA. My interview participant explained the internal logic:

> When we built PISA in 2000, we had the similar effort where we looked at not the measurement, but the conceptualisation of the knowledge, skills, attitudes and values that will be important for people [i.e. the DeSeCo project]. That's […] the kind of the broader picture of the future of education and then we designed PISA… So, there was a process where we actually mapped out those knowledge and skills and then PISA had always seized where we wanted to be. (OECD Interview #2, 2018)

In recognising this connection, I can then identify the normative basis of PISA, in particular what it values and the way in which it seeks to measure, by examining the underpinnings of DeSeCo. Indeed, a closer scrutiny of DeSeCo synthesis reports and strategy papers (OECD 2002, 2003b, 2005b) reveals a train of imperialist assumptions that underlies its approach. For example, the conceptual framework, elaborated in the final report *Key Competencies for a Successful Life and a Well-Functioning Society* (2003b, 4), had been developed based on the context of OECD countries but claimed universal applicability:

> Although DeSeCo was undertaken in the context of OECD countries, there are compelling reasons to assume that DeSeCo's conceptual work and, in particular, the three-fold categorization – interacting in socially heterogeneous groups, acting autonomously, and using tools interactively – constitute a useful conceptual tool for developing and transition countries.

The statement above is just one explicit illustration. In fact, the assumption of DeSeCo's scientific universalism is rather diffuse and very much unified so that it remained largely unnoticed (perhaps even the educational specialists themselves were unselfconscious about their single reading of education and society; see also Silova and Auld 2019).

Given that the key competencies identified within DeSeCo are perceived as 'necessary' for an overall successful life and a well-functioning society, the implication is that those who do not possess these competencies will be categorised (or positioned) as 'not successful' and the society/context within which they are situated as 'ill-functioning'. Consequently, by defining and creating *standards* for what a successful life/person and a healthy society are like, DeSeCo has already succeeded in establishing the normative basis for inclusion/exclusion and, thereby, identity. Figure 4.2 visualises the project's conceptualisation of 'a multi-dimensional space defined by key competencies' (OECD 2002, 15), which I interpret as essentially an imaginary space created through narrowly defined constructs and used to incorporate various contexts.

Figure 4.2 Key competences for an overall successful life and well-functioning society in different contexts.
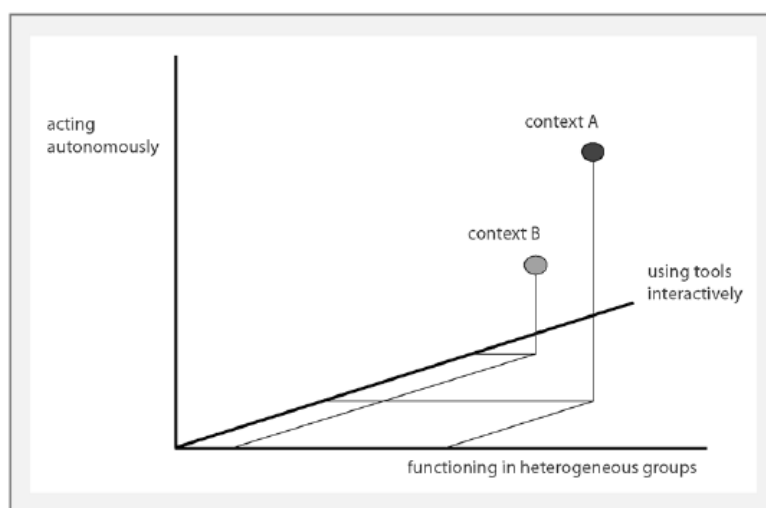
Moreover, and tactically, practices of 'othering' and the creation of hierarchies are also carried out through statistical procedures relating to the *continuous* dimension of competencies. It assumes that 'competencies and their components exist on continua ranging from the low basic level to the high self-authoring level of competence' (ibid., 17). Specifically, this involves the process of

> […] determining where along the continuum from a low to high competence level an individual's performance falls. It is important that frameworks and items in assessments of key competencies include the complete range … By recognizing the existence of continua (or continuous scales, using more technical terms), levels of proficiency (i.e. bands of scores along the scales) can then be identified for *analytic* purposes and for *aiding* interpretation of the scores. (ibid., my emphasis)

As the quote shows, the competence range is not considered to be premised on theoretical underpinnings, such as the three major learning models (reviewed in Chapter 2); rather,

it is designed on a *pragmatic* basis in order to place countries on a common scale and to make the results interpretable. Central to this functioning is the conversion of 'complex social processes and events into simple figures or categories of judgement' (Ball 2003, 217) whereby subordination is enacted in the guise of objective data. I contend that this was the most critical step for the OECD to realise its 'comparative turn', which allowed for the establishment of standardised criteria against which Member countries and non-member countries can be compared and contrasted and located within continuous scales of rankings and ratings. As Cowen (2014, 291) rightly points out, 'comparative education (of this kind) mutates into a system of governance'. The process was greatly facilitated by the OECD's expert authority which is, however, riddled with assumptions, imaginings, and judgements.

Accordingly, PISA is increasingly being marketed by the OECD through media channels (Grey and Morris, 2018) as the most reliable instrument for generating large-scale and comparative data in educational outcomes (see Chapter 5). Starting with 28 Member countries and 4 non-member countries that participated in PISA in 2000, the reach of PISA had gone far beyond the OECD membership to include an additional 30 non-member countries and economies by 2012. With the sheer volume of participants and its distinctive approach of ratings and rankings that generates public scrutiny and peer pressure, the OECD has become an influential actor in *shaping* education policy at both the national and international levels (Nóvoa and Yariv-Mashal 2003; Rinne et al. 2004; Lingard, Rawolle and Taylor 2005; Ertl 2006; Takayama 2008; Grek 2009; Ozga et al. 2011; Benavot and Meyer 2013; Waldow, Takayama and Sung 2014; Morris 2012, 2015; Rubenson 2015; Wiseman and Taylor 2017; Niemann and Martens 2018). Henry et al. (2001, 90) summarise the process of change in the OECD's approach to education, which had moved

> from philosophical doubt to statistical confidence; from covering some countries to covering most of the world; from a focus on inputs to a focus on outputs; and from occupying an experiential status to being a central part of the Organisation's work.

From the constructivist IR perspective, the OECD has evolved from being primarily the instrument of powerful member states (evident in the last period) into an autonomous and influential source of power with its own interests. The OECD's changing role in global educational governance can be attributed to its highly evolved capacity to respond to external changes *and* shape the niche it occupies by both sustaining and extending its sphere of operation. Key to its approach in exerting influence is the technical-rational

authority that the OECD embodies, and through which it creates norms, establishes standards and manages knowledge. Within these processes, the OECD also maintains hierarchies primarily through statistical 'othering' and (re)constructing identity; these aspects are particularly relevant to understanding the OECD's treatment of low- and middle-income countries in the context of PISA-D as it endeavours to extend PISA *globally*. Indeed, Sellar and Lingard (2014) note the OECD's intention to expand its comparisons in education in which PISA has effectively served as the prototype. They suggest that such expansion unfolds in three major domains: the increasing scale of coverage of countries and economies; the broadening scope of what is measured; and the efforts to enhance its explanatory power to inform policies.

However, the more PISA expands, the more limited it seems to be for serving all these purposes and developments. As my interview participant elaborated,

> [s]ince the year 2000, we have expanded PISA by about 10 countries every cycle. When the first countries were included, it was quite easy to accommodate, because they were quite similar in the level of social economic development. But then when PISA grew further, we found actually that we were not doing the best service […] So we felt that at that point we cannot just stretch PISA further and further; we really needed to develop some new tests, some new assessments to use; we needed to redesign the background questionnaires and so on […] Around 2012, we had this kind of first thought that we could not possibly expand PISA in an unlimited way […]. (OECD Interview #2, 2018)

In recognising the limits of the existing PISA framework and instruments for a much more diverse set of countries (and economies), whilst also securing its continual expansion and legitimacy, the OECD once again decided to change gears. The statement above thus has important implications for identifying the features of the OECD's approach to education in the next period, which I have situated within the context of the organisation's transformation for the period post-2011 and the UN's SDGs (2015–2030).

From 2011 to 2030: Sustainable Development Goals and a global learning metric

The OECD celebrated its 50th anniversary in 2011. The Ministerial Council Meeting considered this year as a milestone to take stock of its achievements and to look ahead to frame priorities for the next 50 years. Looking forward was concerned with how the OECD can best respond to global challenges in a rapidly changing world. One of the major challenges identified was the changing global economic landscape in which the economic centre of gravity is shifting from industrialised countries to the large number

of developing and emerging economies. This changing dynamic is clearly illustrated in the following extract from the *Secretary-General's Strategic Orientations for 2011 and beyond*:

> 50 years ago, OECD countries accounted for some 70% of global GDP. Today their share has fallen to about 60% and is set to fall further. … Emerging economies have achieved growth and development perspectives by harnessing market forces to lift their populations out of poverty, just as advanced economies did in their history during the process of industrialisation. Together we can press forward with this process so that markets continue to expand the wellbeing of people in all our countries. Sharing our experiences can guide and unite us in achieving our economic and social goals. This has implications for policy making, for the global governance architecture, and indeed for the OECD itself. (OECD 2011d, 4–5)

In response to the rise of emerging economies leading to global shifts in growth, wealth and influence, the OECD has continued to evolve to position itself as a *global* policy sharing organisation (and/or network). The purpose of this shift in positioning is twofold: first, it allows OECD policy activities not to be confined to its membership so as to incorporate a greater number of developing countries and economies; and second, doing so would ensure greater inclusiveness and continued *relevance* of the OECD's policy work (in, for example, knowledge sharing and mutual accountability) and thereby legitimate and strengthen its global governance potential. In a fast-forwarded reading, these dimensions of the OECD's struggle for survival, adaption and expansion were elicited when Schleicher was interviewed by Addey in 2016:

> […] I think that the OECD membership is somewhat of a historical relic, the OECD has really become a global organization. It shares its global expertise with *whoever* wants to be part of it. It is nothing to do with your wealth or regional location. So basically, it is a *global network of expertise*. […] the alternative is totally irrelevant, if the OECD remains a club of few countries, the reality is that these countries are shrinking in their impact on the world and other countries are rising. I think if the OECD wants to remain global, it has to. (2016a, para.3, my emphasis)

Recognising the need for its work to be more inclusive and responsive to the changing global realities, the OECD, therefore, laid out an all-encompassing *Framework for an OECD Strategy on Development* (2011f; adopted by the OECD Council in January 2012). The Framework outlines a broader development strategy for the OECD, which seeks to synthesise its 'multidisciplinary expertise on a wider range of policy areas with its accumulated knowledge on development issues, and to ensure that this work is accessible and relevant to countries at all stages of development' (OECD 2011a, 3). The Framework also responds to the demand for governance reform in both developed and developing countries, fuelled by the global crises of recent years – financial, economic, food, and energy crises and climate change (OECD 2011f).

In my analysis of the second period (from the 1990s to 2010), I noted that there had been a shift in the governance approach within OECD countries from micromanaging to central steering, which was greatly facilitated by the OECD setting the overall framework. In the post-2011 context, however, a new approach to governance seems to emerge. This new approach is informed by the OECD Framework (noted above) and oriented towards on a 'more coherent and comprehensive approach to development' (2011a, 3) as the OECD seeks to expand its policy instruments into low- and middle-income countries. The following extract from the key OECD document, *Better Policies for Development: Report on the DevGoals Exercise*, supports this point:

> […] the OECD has made significant progress in bringing its development experience together with its multidisciplinary expertise to build and apply a broad approach to development … This is a reflection of the external demand for - and validation of - the OECD's comprehensive approach which looks at development beyond aid, is growth-oriented and based on knowledge sharing and policy dialogue.

> The challenges now are to apply this broad approach *across* the organisation, exploit the synergies more systematically among these strands of expertise within the Organisation; enhance OECD's working methods; and *expand* policy sharing *beyond* the membership. The OECD's 50th anniversary provides an opportunity to reinforce the Organisation's commitment to development and move towards a new OECD Strategy for Development. One that: combines effectively OECD's collective expertise; integrates the demands of developing countries; strengthens partnerships with other international organisations; ensures policy coherence for development; and reinforces OECD's working methods. (2011a, 4, my emphasis)

Elaborated in the 2012 strategy paper, *OECD Strategy on Development*, this broad approach to development comprises four thematic areas where the OECD is positioned to 'have core competence' and 'could add value to other international efforts, respond to the demands and needs of developing countries, and leverages its multidisciplinary expertise' (2012d, 5). The four areas are (i) innovative and sustainable sources of growth, (ii) mobilising resources for development, (iii) governance for development, and (iv) measuring progress for development. Taken together,

> they offer a holistic framework from which member and partner countries could identify a set of specific instruments, mechanisms and tools in interconnected policy areas [of the economic, social and environmental], relevant to their particular circumstances and needs. (ibid.)

As these extracts indicate, this emerging approach is distinctive from the previous approach to governance within OECD countries that functioned through translating complex processes and events into simple categories and figures to establish *standardised* criteria (e.g. the INES project and DeSeCo). The emerging mechanism is one in which PISA is part of a 'configuration', as the OECD staff member described, that provides

multidisciplinary expertise and a governance toolkit, which can be *tailored to* countries at all stages of development. Rather than as being a 'one-size-fit-all' approach, this emerging approach is more inclusive *and* flexible. To use Hardt and Negri's (2001, 198) metaphor, it operates like 'a machine for universal integration, an open mouth with infinite appetite, inviting all to come peacefully within its domain'. I thus interpret this emerging approach as representing the formation of an advanced form of governance mechanism, which sustains the OECD's central steering while allowing it to constantly expand the realm of its control to the furthest corner of the globe (explained further below).

To implement this comprehensive governance approach, the OECD identified three complementary levels of engagement that it needs to strengthen (OECD 2012a, 2012b, 2012d). The first level of engagement is working with OECD *Member countries* to support their collective efforts to design policies consistent with development. The primary objective in this regard is to enhance policy coherence, based on more systematic approaches to evidence-based analyses and robust indicators to monitor progress and assess their impact on development. The second level of engagement is working with other *international organisations* and *international efforts* to seek solutions to address current and emerging global issues and development challenges. The major concerns in this respect are the Sustainable Development Goals (SDGs) and the effectiveness of development co-operation highlighted at the Busan High-Level Forum (2011b). The third level of engagement is working with *developing countries* to promote knowledge sharing and policy dialogue activities, such as multi-dimensional country reviews and the clustering of countries with similar economic challenges.

Within this approach, education was identified as a cross-cutting theme reflecting a series of related development efforts (i.e. *Education for Development*). For example, it noted that 'progress in education, training and human resource development is essential for achieving sustained economic growth and poverty reduction, and for responding to the changes in technology and demographics that are re-shaping labour markets', and 'education systems need to provide high quality and equitable learning opportunities – starting in early childhood and continuing throughout life'. Moreover, 'people with sound baseline qualifications who then enter into targeted training program systems are able to achieve higher skill levels and possibly increased mobility' (OECD 2013b, 53).

In terms of educational action plans outlined under the theme of *Education for Development*, it was suggested that enhancing learning outcomes and making them visible remained crucial. The imperative was related to the series of high-level fora on aid-effectiveness (OECD 2003a, 2005a, 2008, 2011b), which had been held to assess progress on the MDGs. A broad consensus emerged during this period among influential donors that investments in education had not yielded proportionate gains in outcomes, followed by the call for greater accountability through the identification of basic minimum standards of education quality. As will be further demonstrated in Chapter 5, the OECD was influential in shaping and promoting this discourse through engaging with UNESCO while it was framing the education SDG (i.e. SDG 4). The process was described as follows:

> There was a big conference in Korea by UNESCO where actually they were framing the SDGs. A big transition is that the past MDGs was about access then focus came to quality. Quality has always been the OECD paradigm, but became a more UNESCO paradigm. But we got involved in the UNESCO set-up Task Forces. […] so we are very actively taking part in UNESCO's work to help UNESCO develop metrics, frameworks and tools to actually bring the SDG 4 Goal to life. Because, again, without those tools, that is just words and I think UNESCO has been using our expertise and our instruments to make the SDGs measurable. (OECD interview #2, 2018)

The expertise and instruments that UNESCO has drawn on to inform and frame SDG 4 largely refer to PISA – a key aspect of the OECD's competence and source of authority. The quality paradigm reflected in the OECD approach works well with the intention to extend PISA into developing countries and to relate PISA to development – PISA for Development. These are clearly illustrated in the *Progress Report on the Implementation of the OECD Strategy on Development* (2013b):

> [1] While the education-related MDGs focused primarily on access to primary education, a consensus is emerging that we also need to improve the quality of education, to achieve better learning outcomes. PISA provides a comprehensive and rigorous international assessment of learning outcomes and, as such, is well suited as a metric to measure learning performance in a post-2015 framework.

> [2] Over 70 countries and economies have benefitted from participating in […PISA]. PISA has given them the opportunity to benchmark progress, learn from best practice and sharpen the focus of their policies on key learning outcomes. … The OECD plans to extend these benefits to a greater number of developing countries including both middle-income (MIC) and least-developed (LDC) countries. A new project will explore how PISA survey instruments can be enhanced to be more relevant for the contexts found in developing countries, while producing scores that are on the same scale as the main PISA assessment. In a post-MDG context, broadening participation in PISA will also provide an opportunity for the international community to measure progress in improving learning outcomes. (37–53)

When situated within the OECD's strategy for development (which endorses a more comprehensive approach to development and moves towards a global policy-sharing organisation), PISA-D can be understood as in strategic alignment with the OECD's broader vision for 2011 and beyond. Amidst the OECD's configuration of policy instruments, it was identified that 'some have a potentially global reach, such as the generation and analysis of data and indicators on financing for development or measuring progress of societies' (OECD 2012a, 34). Shortly after this background report for the strategy was laid out, the Education Committee and the DAC 'started to reflect on how they could better capitalise on the PISA tool and make it more applicable for a wider range of developing countries' (OECD 2013b, 12). It was against this background as the OECD seeks to leverage its multidisciplinary expertise by scaling up selected activities with global reach that the idea of 'PISA for Development' was born. This also aligns with the other two levels of engagement envisaged in the Framework: with other IOs and the SDGs, and with developing countries. The OECD internal (re)structuring dynamics, however, contradict the OECD portrayal that PISA-D was introduced primarily in response to the demand of low- and middle-income countries (see Chapter 5).

The significance of PISA-D was soon acknowledged in the *2014 Report on the Implementation of the OECD Strategy on Development*. I provide the extract below:

> PISA for Development is well placed to support global efforts to frame a learning goal in the context of the post-2015 agenda and to provide a single universal metric for measuring progress on this. PISA is regarded as one of the most important education policy instruments in the world today. There is a compelling logic for making this more relevant to developing countries so that greater numbers can benefit from the surveys and analysis. The project therefore addresses *squarely* the *core elements* of the Strategy on Development, particularly the *adaptation* of *the most successful OECD policy instruments* to make these more relevant for *developing country contexts*. (2014a, 10, my emphasis)

When the PISA-D project was completed and declared successful in September 2019, the strategic alignment efforts noted above were revisited in the *Project Completion Report* (OECD 2019c). This report celebrated all countries' commitment to the universal education agenda, stating that for the OECD, 'this means an acceleration of the transition envisaged in the 2012 Strategy on Development towards more inclusive policy dialogue and the participation of a wider array of countries in its activities' (ibid., 89). A noticeable development is that PISA 'is [now] the world's most *inclusive* metric on student outcomes and one of the most *global* instruments of the OECD', and its significance was stated as 'not just for monitoring SDG progress but also for informing policy and funding decisions

that will help countries *achieve* the SDGs' (ibid., 90, my emphasis; see Chapter 5 for an analysis of the legitimation strategies). Thus, through PISA-D, the OECD has succeeded in including more non-Members, especially low- and middle-income countries, in its policy network. By facilitating peer-to-peer learning among these new and existing participants, the OECD can also achieve greater inclusiveness and utility of its policy expertise, and is, therefore, better positioned to shape the future global architecture of development.

In the report entitled *Looking Ahead to Global Development Beyond 2015: Lessons Learnt from the Initial Implementation Phase of the OECD Strategy on Development* (2014d), the OECD is portrayed as having achieved success in the initial implementation phase of its development strategy. It also stated that the strategy 'is evolutionary in nature, recognising the need for the Organisation's work on development to be more responsive to complex global realities' (OECD 2014d, 2). In the current post-2015 context defined by the SDGs, the OECD continues to frame its work to align with inclusive and sustainable development discourses. It underscores that strengthening policy dialogue and knowledge sharing is critical to adapting its analytical frameworks and policy tools, and to respond better to the contemporary world that requires multidimensional solutions.

The adaptation of the analytical frameworks is now taking shape in the OECD Learning Framework 2030, which aims to update DeSeCo, to identify future knowledge, skills, attitudes and values that today's students will need to acquire to succeed in the 21st century (OECD 2018a, 2019a). Auld and Morris (2019, 2) suggest that the OECD 2018 assessment of 'global competency' is part of this broader strategy to 'extend beyond the cognitive-economic focus of "education quality" to the non-cognitive dimensions highlighted in SDG 4.7', namely

> […] 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. (UNESCO 2015, 73)

However, they demonstrate how the measurement of global competencies was, in contrast to PISA-D, retrofitted at a late stage of its development to link to the SDGs.

While UNESCO's decision to identify basic minimum standards of education quality under SDG 4.1 (2015–2030) marked the shift from Education for All (EFA) to learning for all, the OECD's vision for 2011 and beyond can be better understood as *assessment for all*. The ensuing call for a 'data revolution' in developing countries made in the OECD

Development Co-operation Report 2017: *Data for Development* constitutes the point of arrival of this tendency. The report underlines the lack of statistical capacity in most developing countries as the biggest obstacle for achieving the SDGs, and it suggests greater investment in (digital) data infrastructure and the enforcement of legal, ethical and quality standards. The following extract from the Secretary-General's remarks at this year's report launch in London provides an illustration of this problem:

> Globally, we cannot expect to deliver on the SDGs with gaps in basic data and weak statistical systems in developing countries … New tools, and the so-called "data revolution" makes it easier, faster and cheaper to produce and analyse data from a wide range of sources … I have often said that when it comes to the SDGs, the international community needs a "GPS". We need to know where we are, where we are going, and what the best route to get there is. This is the power of good data. Data provision and analysis is in the OECD's DNA. It goes to the heart of all the areas in which we work, including development. (Gurría 2017, para. 14)

These remarks imply that the OECD's focus on data provision and analysis are self-serving activities. Their existence mirrors what Hardt and Negri (2001) call the 'one machine' which thrives by always including more within its sphere (if there is no data infrastructure, we will help to build it). PISA-D is the latest manifestation of this endeavour, which has helped to establish some initial data infrastructure in low- and middle-income countries for PISA data collection and analysis. Chapters 5 and 6 will show in more detail how these activities were carried out through 'capacity building' events. Therefore, the cloak of PISA *for development*, or data *for development*, prepares for universal acceptance of *assessment for all*. We are currently witnessing the OECD's departure from the 'comparative turn' in the late 1990s and its new 'turn' to take on a 'humanitarian' approach which has been facilitated through engaging with the SDGs and aligning PISA data with the monitoring efforts. As the OECD puts it, "[w]ith the shift away from government "of the people" to government "for the people", data needs have changed' (OECD 2017a, 51).

I suggest that the OECD's governance approach is currently going through a steady and gradual transition. While in the previous period it focused on establishing standardised criteria and setting the overall framework, 'configuration' and modulation are now at the core of the OECD strategy. Specifically, the emerging approach operates through a 'configuration' of the OECD's multi-disciplinary expertise and a governance toolkit, which can be optimally applied to countries across all geographies and at all stages of development. Modulation refers to the process of tailoring that coordinates differences under a hierarchical command structure. To illustrate these points further,

cross-referring of the following extract from my interview with the OECD staff member on PISA-D post-pilot governance is again helpful:

> We are now integrating the PISA-D instruments into PISA so that every country can actually tailor PISA to its national needs. PISA in the past was one instrument for everybody and now it becomes a *configuration* that countries can actually *tailor to* their needs. In the future, actually we will make PISA *individually adaptive* … (OECD Interview #2, 2018, my emphasis)

The specific plan for integration and the future use of PISA-D instruments has been outlined in the *Project Completion Report* (2019b). It states that the OECD is seeking to build on the success of PISA-D to continue to make PISA more accessible and relevant to a wider range of countries. In order to cope effectively with the demand from different participants, the plan has put into place 'next steps' for managing 'a larger PISA programme' (ibid., 102) which I consider as the 'configuration' that my interview participant described. To this end, a total of 13 key steps have been listed in the report, which are summarised in Table 4.2.

Whereas old power seeks to fix pure, separate identities, the new one thrives on 'circuits of movement and mixture' (Hardt and Negri 2001, 199). In the case of the OECD, its previous apparatus was a kind of mould that fixed distinct castings (e.g. PISA as one standardised instrument for everybody), but the emerging one is a 'configuration', as my interview participant highlighted, that can be tailored to the needs of individual countries regardless of their levels of development. Configuration in the sense also that the programme builds on a universal foundation of similitude that applies equally to all countries, echoing d'Agnese's (2015, 56) point about the OECD imposing its own 'univocal logic'. Similar to the process of capital (in Hardt and Negri's depiction) bringing all forms of value together on one common plane and linking them all through money, the OECD's univocal logic works through objective data. As the world is moving into 'an educational intelligent economy' premised on the exponential production of digital data (Salajan and Jules 2020, 1), it seems to bear some resemblance.

Therefore, modulation is the process of simultaneously accepting and acknowledging differences and ensuring the 'configuration' continues to make sense (an example of this is the EPA applied to link performance to contextual factors in each PISA-D participating country; see Chapter 6). This process is 'needed to take into account the heterogeneity of growth and development models, and the differentiated institutional settings and capacity to tap into resources and address any binding constraints' (OECD 2012d, 4). Accordingly, the OECD's realm of operation and control are not exercised through binary divisions

(such as the West and the East, or the North and the South), but through a strategy of *differentiated* inclusion. As the OECD (2012a, 10) states, '[t]here are no one-size-fits-all solutions in development'; differentiation – '[a]dopting differentiated approaches, more diagnostic than prescriptive, will be key to identifying the varying needs of countries at different stages of development and develop better analysis considering their diverse economic and social contexts'.

Therefore, at the surface level, it seems that

> [p]olicy prescription is not made through any normative or authoritarian process, but instead allowing each nation state, in its freedom, to do what is "right", to do "what needs to be done", to do "what works". That is, it is not conspiracy, it is an interweaving of free-wills'. (Nóvoa 2017, 6)

However, I suggest that the OECD's dominance and statistical 'othering' are sustained beneath its magnanimous façade. These are practised through first engaging with low- and middle-income countries and then inviting them to follow the path dictated by PISA. In this way, differences are subordinated to the OECD's established technical standards, which are, essentially, not merely technical. 'Commensuration as a practical task requires enormous organization and discipline that has become largely invisible to us' (Espeland and Stevens 1998, 315). Similarly, Pettersson and Popkewitz (2019, 33) argue that 'expertism' is something that has constantly to be seized and debated, and they observe that one of the major impacts of PISA has been to place the expertise in education in the hands of entrepreneurs, technicians, economists and statisticians. Since discretion is replaced by disciplined methods, the latter also diminish autonomy and thus 'real' free-will.

Table 4.2 The OECD's plan for managing a larger PISA programme – *next steps*.

1. **To provide optional support for new LMIC Partners** to help them prepare for their participation in PISA and assist them with survey implementation.

2. **To encourage new LMIC Partners to concentrate in their first cycle on the implementation of PISA's core components** and **will not be offered the alternative to take on additional options** such as optional questionnaires or tests, over-sampling and the optional out-of-school component **until they have the experience of completing at least one cycle** successfully.

3. **To promote peer-to-peer learning partnerships** between **an experienced** PISA country and **a new** country to help new National Centres to participate and complete the necessary phases of project implementation.

4. **To enhance the ways results are presented** in the PISA international reports to reflect the wider participation in the programme and to make available the full range of data.

5. **To continue to support analysis and reporting** for LMIC partners in PISA 2021 and PISA 2024.

6. **To discuss support for countries to prepare regional reports.**

7. **To maintain the use of paper-based assessment** in PISA 2021 and PISA 2024 and **to investigate the comparability of paper- and computer-based scales**.

8. **To develop technical review of the PISA scaling methods to accommodate further non-Member participation**.

9. **To continue to enhance PISA instruments to better describe the performance and contexts of a wider range of students**.

10. **To extend the contextual questionnaires in PISA 2021 to measure economic, social and cultural status** (ESCS) to better describe the contexts of a broader variety of socio-economic contexts in **all** Partner and Member countries.

11. **To explore how the PISA-D out-of-school component can be integrated in future PISA cycles** as an optional module.

12. **To provide the incorporation of out-of-school component as an optional module** so countries who choose this option can obtain information about the skills acquired by all children.

13. **To streamline and manage PISA NPMs** (National Project Managers) **meetings** considering the increased number of participants. The streamlining measures include **limiting** the number of country representatives at each meeting; adding **content-specific** days to NPM meetings; holding special NPM meetings for new countries; and improving NPM governance.

Source: OECD (2019b, 102–103)

Summary

In this chapter, through interpreting the three distinctive time periods in the OECD's history, I have shown that its agendas and approaches to education have evolved to make its work more responsive to changing global realities, which in turn have shaped perceptions of those realities and the ways in which they can be addressed. I have further shown that the OECD's history is one of successful expansion. Thus, the OECD's approach to education is also expansionist in nature, as it continues to innovate, find market niches and add new dimensions to its data sets. These have culminated in the work of PISA-D, which strategically moves forward the OECD's strategy for engaging with developing countries and the SDGs.

The periodisation also shows that the origins of PISA-D lie in the Cold War and the influence of American power which gave rise to its ideological basis. Its prototype, PISA, which was developed in the late 1990s, was integral to the OECD 'comparative turn' for developing cross-national comparisons (i.e. INES) and formulating an overarching assessment framework (i.e. DeSeCo). Accordingly, PISA-D is the continuation of the OECD's endeavour to reinvent itself and to increase its power and relevance.

In the post-2015 era, and with the declared success of PISA-D, it would appear that the OECD has successfully manoeuvered into a position whereby it can continue to shape the future of education and development. The current period is distinctive in that the OECD claims to take on a comprehensive approach to development, integrating social and environmental concerns into the economics of education approach. It suggests that a unidimensional approach to addressing economic issues misses many of the systematic interactions that really define the system. Thus, with the advancement of machine learning and big data, it is moving towards configurating multidisciplinary expertise and a governance toolkit to transform and upgrade its capabilities.

Consequently, education has been recognised as a cross-cutting theme within its policy agenda, and the use of data – the OECD's DNA – is infused into monitoring the related progress and impact on development. The proposal for a 'data revolution' in low- and middle-income countries (OECD 2017a) and the plan for composing 'a larger PISA programme' (OECD 2019c) illustrate well this tendency. From these perspectives, the work of PISA-D is significant for accelerating this transformation. The OECD is well on its way towards achieving assessment for all.

However, as shown in Chapter 2, the evidence is not sufficient to demonstrate that PISA is contributory to development. PISA has not yet taught us much about learning or has improved educational outcomes (Baird et al. 2017; Kanjee 2018), and it is unlikely to bring about higher GDP growth rates (Rappleye and Kamatsu 2017, 2019; Feniger and Atia 2019; Patel and Sandefur 2020). These point towards my next argument that the meaning of 'PISA for development' is being contested and skilfully managed. It involves what Suchman (1995, 579) refers to as 'self-serving claims' with 'hollow symbolic gestures'. What is critical about these claims is that they are persuasive, and hence effective in bringing a form of moral authority to the OECD for monitoring SDG 4. Legitimacy management is also essential especially when the OECD finds that PISA continues to fail to address field realities, which could threaten its scientific credentials. These dynamics are important to deepen our understanding of the OECD's operating logic of continual expansion and organisational survival. The following two chapters elaborate more on these aspects.

Chapter 5. PISA-D Legitimation Strategies

An international organisation's political authority is at its zenith when its rational/technical agenda aligns with prevailing social values and sentiments.

Richard Eccleston, *The OECD and Global Economic Governance*

We have an obligation to try to make this work for the sake of development and to support the enhancement of global skills levels.

Andreas Schleicher, presentation at the *Initial PISA-D Technical Meeting*

Introduction

In this chapter, I identify the major strategies deployed by the OECD to legitimate PISA-D. This is pursued through a critical hermeneutic approach to OECD documents, focusing on the 'interrelated set of texts, and the practices of their production, dissemination, and reception', which Phillips and Hardy (2002, 3) broadly define as 'discourse', that renders PISA-D into a legitimate object. In Chapter 4, I proposed that PISA-D emerged through the progression of PISA as the OECD sought to leverage its global expertise and governance toolkit and to extend them into low- and middle-income countries. To this end, the OECD has actively engaged with post-2015 discussions about the formation of the SDGs and has been increasingly vocal about bringing its own expertise (in particular, its testing instruments) to SDG 4 monitoring. The focal point of this chapter is to trace how the OECD has accomplished these, while overcoming the existing critiques of PISA for low- and middle-income countries, and strategically moved forward its transformation for 2011 and beyond.

Overall, I identified six strategies to be critical: (1) *building affiliations*; (2) *shaping and (re)framing SDG 4*; (3) *establishing credibility*; (4) *creating demand*; (5) *membership and belonging*; and (6) *persuasion in motion*. Table 5.1 summarises the major practices and processes of each strategy. As previously mentioned in the introductory chapter, some of these strategies bear resemblance to marketing techniques, especially those elements pertaining to demand creation and membership and belonging. However, below the surface of these techniques, I also view that the OECD continues to hold imperialist assumptions of superiority and otherness, echoing Silova and Auld's (2019) depiction of 'banal imperialism'. Although Silova and Auld suggest that the OECD's world order of the juxtaposition of the East and the West (or the South and the North) is 'breached' as certain non-Western countries outperform the OECD countries, my reading is that such a

'breach' can be the result of a deliberate 'circuit of movement and mixture' (Hardt and Negri 2001), characterised by inclusion and flexibility, which the OECD applies to expand its imperial realm of control.

Accordingly, in the case of the legitimation of PISA-D, I will illuminate the ways in which the OECD 'peacefully' integrates low- and middle-income countries into a *global* PISA community, whereby the other is elaborated through degrees of deviation relative to the OECD average and/or top-performing or fast-improving countries. Within this realm of control, I contend that while certain countries (e.g. Brazil, Estonia or Kazakhstan) may creep up in the ranks (and are appraised), the order itself – the PISA ladder – remains rather still. Consequently, countries, especially low- and middle-income ones, are always defined in terms of a *place* within the OECD grid as, for instance, Figure 4.2 shows. The related practices (or specificities) of such an operation will be discussed further in Chapter 6 when I examine the PISA-D outcomes. In the following sections, I provide details of these legitimation strategies with case materials selected from the OECD documents, and accompanied by cross-referencing to interviews and literature.

Table 5.1 Summary of PISA-D legitimation strategies.

5.1 Building affiliations	(a) Affiliating with UN's specialised agencies in education and development: UNESCO & UNICEF.
	(b) Elaborating on its experience and expertise (in PISA and related policy analysis) as valuable contributions to SDG 4 formation.
	(c) Positing itself as an integral part of the decision-making process in framing SDG 4 and utilising this position to improve sociopolitical acceptability of PISA and PISA-D as the learning metric.
	(d) Taking on a 'collaborative turn' to highlight established partnerships with these agencies to help raise the profile of PISA and PISA-D as the prominent metric for monitoring SDG 4 targets.
5.2 Shaping and (re)framing SDG 4	(a) Narrowing the focus of SDG 4 from all stages to only secondary education and limiting the scope of educational outcomes to what is most relevant to PISA measures.
	(b) Highlighting the critical role of Indicator 4.1.1(c) in achieving SDG 4 and leveraging the use of PISA and PISA-D as the optimal solution to monitoring SDG 4 globally.
5.3 Establishing credibility	(a) Confirming PISA's established record of credibility by stressing (i) the collaborative nature of its work, (ii) the validity and reliability of the testing instruments, and (iii) the global scale of its coverage.
	(b) Utilising exemplary cases (such as Brazil, Peru and Vietnam) to demonstrate the benefits of PISA for low- and middle-income nations (thereby encouraging their participation in PISA-D).
5.4 Creating demand	(a) Shifting from the 'supply-side' story to the demand-led: PISA-D was introduced primarily in response to (i) the demand low- and middle-income countries who wanted to participate in PISA, and (ii) the needs of the international community for a universal learning metric on SDG 4.
	(b) Involving 'a self-perpetuating dynamic' to consolidate the status of PISA and its augmentation by PISA-D.
5.5 Membership and belonging	(a) Establishing a strong sense of membership and belonging to the global PISA community.
	(b) Members of the community enjoy such benefits as: (i) international benchmarking and comparison; (ii) access to technical expertise; (iii) good practices and peer-learning; and (iv) demonstration of commitment.
5.6 Persuasion in motion	(a) Holding meetings, workshops and seminars on a global scale and ensuring its completion and 'success'.
	(b) Creating the microspaces (i.e. social settings and events of speaking and exchange) whereby PISA-D discourses continued to circulate and gained momentum.

5.1 Building affiliations

According to Edwards et al. (2018, 34), 'an organization can dedicate itself to achieving sociopolitical acceptability by *(a)* adapting to its context, *(b)* responding to changing expectations, and *(c)* invoking or affiliating itself with symbols (or other organizations) that possess legitimacy'. In this study, the first two dimensions relating to the OECD's adaptation and responsiveness are illustrated in Chapter 4. Here, the third aspect about invoking (or affiliating with) other organisations that possess authority resonates with the strategy the OECD took *at the outset* to legitimate PISA-D.

As I have noted, while the OECD was previously involved in the development of the MDGs, that work was primarily concerned with aid coordination and effectiveness through the DAC (OECD 1996a). Moreover, at the time the MDGs were formulated in 2000, the OECD's flagship programme, PISA, was still in its infancy. Therefore, the OECD was keen to play a major role in post-2015 discussions on the SDGs. This would allow for 'a more comprehensive OECD contribution and increased relevance, value added and impact in major international processes' (OECD 2012d, 6).

To this aim, the OECD started by invoking or affiliating itself with the UN agencies that possess legitimacy to frame the SDGs (in particular SDG 4). This endeavour and tone are most clearly illustrated in the extract below from the 2013 *Progress Report on the Implementation of the OECD Strategy on Development*:

> [1] The OECD is committed to supporting a post-2015 agenda, and to this end is closely *following*, and *where invited*, contributing to discussions at the UN on how to frame the agenda. For example, the Education Directorate is an active member of the Advisory Group on the post-2015 education agenda, co-chaired by UNESCO and UNICEF. Most recently, EDU helped moderate an online discussion on the quality of education. (2013b, 24–25, my emphasis)

> [2] There was a big conference in Korea by UNESCO where actually they were framing the SDGs. A big transition in the past MDGs was about access, the focus came to quality. So that's basically the UNESCO discussion, which the OECD wasn't involved in. (OECD Interview #2, 2018)

Although we already know retrospectively that the major outcome of these engagements is PISA-D, to *affiliate* itself *with* these key agencies and to *gain access* to discussions at the UN was the first and foremost result for the OECD at such an early stage. Once the OECD joins the discussions, then it can be increasingly vocal about its vision and approach to education and development (i.e. shaping perceptions of educational realities, as one of the OECD staff members in Chapter 4 stated; see also the shift in the OECD's discourse in the following extracts).

Accordingly, given that UNESCO and UNICEF are the officially recognised custodians of SDG 4 indicators within the UN system, the OECD posited early on that it was critical to relate to and to collaborate with these agencies. This can be observed in the extract from the *2014 Report on the Implementation of the OECD Strategy on Development*, where it identified that working in close collaboration with these lead agencies is a critical factor for PISA-D success:

> Critical factors for [PISA-D] success include: engaging at the *right* level, *politically* and technically; ensuring a robust design of the concept, relevant to partners' interests and needs; working in collaboration with development partners to approach developing countries, avoiding duplication of donor country government agreements and arrangements for working in education; and *working in close collaboration with international agencies that have a mandate on global education,* such as *UNESCO* and *UNICEF*. (2014a, 7, my emphasis)

From a constructivist IR point of view, as bureaucracies, IOs are conferred rational-legal authority, but they also derive authority from delegation processes, moral claims, and expertise (Barnett and Finnemore 2004). In this case, what the OECD sought was a sort of moral and delegated authority to further its political agenda. Moral claims, aligned with its expertise (see below), help its voice to be heard, recognised, and believed, whereas considering the lack of delegated/mandated responsibility in framing the SDGs, the strategy to affiliate itself with and assist these key agencies helps to enhance its sociopolitical acceptability.

The question is then how such an affiliation was established. To put it differently, how did the OECD relate to and collaborate with these agencies? To illustrate this point, it is helpful to first look at some extracts from the OECD's *Post-2015 Reflection Series*:

> [1] UNESCO has established a Learning Metric Task Force with the support of the Brookings Institute to bring together *experts* to develop recommendations for the global education and development communities about internationally comparable learning standards, metrics and implementation practices... The OECD **will** *contribute experience, evidence, analysis* and *relevant policy knowledge to* the Learning Metric Task Force, the UN's thematic consultation and other relevant fora that are responsible for developing and agreeing post-2015 education goals.

> [2] …the OECD **will** *support* education-related post-2015 goals […]. In particular, OECD **will** engage in the relevant post-2015 forums convened under the UN and *contribute to* these *on the basis of* its *extensive experience* in using evidence emerging from *the analysis of PISA data,* complemented by insights developed by its *expertise in policy analysis, implementation and research and innovation*, **to inform** the development of a new generation of global goals and targets in education. (2013c, 6–7, my emphasis)

As the extracts show, the OECD's collaboration has to do with its expertise gained through years of developing PISA (although as discussed in Chapter 4, this kind of 'expertism' is not given but has been constantly constructed; see also the 3rd strategy

identified below). Specifically, the OECD invoked its merits by elaborating on its experience and expertise in PISA which can be perceived as *valuable* contributions to the post-2015 agenda. The key to this endeavour was the creation of what Suchman (1995, 574) defines as a 'generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate'. By highlighting its authority of expertise and closely 'collaborating' with these key agencies, the OECD has thus assumed its sociopolitical acceptability. Moreover, what these extracts also show is that the OECD was at the periphery from the outset, while UNESCO and other related UN bodies were framing the education goal and targets; therefore, their legitimate roles and authority were fully acknowledged by the OECD.

Once the affiliation was established and linkages built, the OECD became more affirmative about its contributions to the post-2015 education agenda. This subtle shift can be observed from the following extracts which embody two aspects: first, the OECD came to position itself as an *integral* part of the decision-making process for framing SDG 4 (also evident in the change of tense from the future, shown above, to the present continuous as below); and second, it proceeded to conceive itself as *well placed* to support the framing and *monitoring* of the education targets, based on its substantial experience in PISA (in other words, it paved the way for introducing and legitimating PISA-D):

> [1] A post-2015 learning goal will *present* the international community *with the challenge* to develop an agreed method for measuring progress. The OECD *is well-placed to* contribute to the definition of learning goals and targets, based on the experience of the Programme for International Student Assessment (*PISA*)… The OECD *has much to share from the lessons from PISA* on *how to measure* learning, the likely pace of progress towards achieving a learning goal and also the importance of avoiding setting over-ambitious learning goals and targets. (OECD 2013c, 1–5, my emphasis)

> [2] The OECD *is engaging in* the post-2015 forums convened by UN agencies such as UNESCO and UNICEF. In supporting the post-2015 process, the OECD aims to *help establish* a set of measurable, globally relevant education goals that are focused on the quality of learning, and that *enable all countries* to set meaningful targets that can be tracked over time. The OECD *is well placed* to contribute to this effort, *based on its experience* with the highly collaborative Programme for International Student Assessment *(PISA)*. (Davidson, Ward and Palma OECD 2014, para. 5, my emphasis)

These extracts are illustrative of how the OECD used its expert authority to strategically weave PISA into the post-2015 education architecture, which was formally adopted as SDG 4 by the UN General Assembly in 2015.

With its increased relevance and acceptability, the OECD then made another twist – by trying to highlight the *collaborative* relationships it had established with these key

agencies to help raise the profile of PISA, and thereby PISA-D (a 'collaborative turn'?). A noticeable example of this was the welcome speech given by Michael Ward (who is the PISA-D project manager and a senior policy analyst at the OECD) at the 5th National Project Managers (NPM) Meeting held in Cambodia on 22 May 2017. He stated that

> PISA for Development is a *very important project* for OECD and for the countries, but also *in the context of the Sustainable Development Goals*, particularly the Education Sustainable Goal.... And *we* were talking about how in the *United Nations, UNESCO,* how PISA *has been chosen* as the metric, *the global metric for measuring progress* towards achievement of the Education Sustainable Development Goal. (my emphasis)

It is worth noting that at the time this NPM meeting was held, PISA had *not yet* been officially endorsed by UIS as the measure for SDG 4 (more specifically, for indicator 4.1.1c). It was only recognised in 2019 as *a* source of data for this indicator (see the 2nd strategy identified below). Similarly, when the OECD published the *PISA-D Analytical and Methodological Framework* in 2018, it highlighted its role in working with these agencies in both framing and monitoring SDG 4, and acknowledged the support it had garnered from these agencies for PISA-D:

> [1] The OECD has been *a key partner* of UNESCO and the other co-convening agencies in developing the Education SDG framework, and *works closely with UIS* in the development of indicators that will be used to measure progress towards SDG achievement. *In turn, UNESCO, UIS and the World Bank **have partnered with** the OECD **in support of** the PISA-D initiative.*

> [2] The OECD, UIS and the World Bank ***are working together*** and with other key […] stakeholders committed to improving learning outcomes in all countries – particularly low and middle-income countries. PISA-D and the OECD's plans for mainstreaming the outputs of the project in the future cycles of PISA is *a key contribution* to these efforts, and ***an embodiment of international collaboration*** in support of the measurement and monitoring of learning outcomes in the context of the Education SDG. (OECD 2018b, 22–23, my emphasis)

While the OECD has so far succeeded in claiming a pivotal role in defining SDG 4 and is well positioned to shape the future architecture of education and development (as shown in Chapter 4), it is important to recognise and trace how it had managed to create its niche and leverage its position. This was pursued at the beginning through invoking or affiliating itself with key UN agencies that possess legitimacy in framing and monitoring SDG 4. It was through this first strategy that the OECD joined the 'high table', and then proceeded to take on an active role in shaping the priorities of the post-2015 education agenda and the way in which they can be best achieved:

> My view on this is that SDG 4 will not have meaning without our capacity *to monitor* progress. You know, we can talk about quality education, equity education. There were lots of words around it, what we cannot measure we will never improve, and that is why I think those kinds of metrics

are really important and PISA gives countries a framework to measure progress. (OECD Interview #2, 2018, my emphasis)

I demonstrate below the processes by which the OECD steered post-2015 discussions towards the legitimation of PISA and PISA-D.

5.2 Shaping and (re)framing SDG 4

The global education agenda encapsulated in SDG 4 expresses a vision for education that is far broader and inclusive than the EFA which accompanied the MDG framework (Unterhalter 2019). That means, as summarised in Table 5.2, SDG 4 comprises targets for expanding access to all stages of education – pre-primary, primary, secondary, vocational, higher and adult education. The targets also recognise a wider scope of educational outcomes, such as global citizenship, sustainability and gender equality. However, for the OECD to promote PISA and PISA-D as the legitimate learning metric, it would have to shape and (re)frame the education agenda in such a way that PISA can be of strategic relevance.

Table 5.2 SDG 4: goal and targets.

SDG 4: one goal, ten targets
SDG 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all by 2030
Targets: **4.1** Universal primary and secondary education **4.2** Early childhood development and universal pre-primary education **4.3** Equal access to technical, vocational and tertiary education **4.4** Relevant skills for decent work **4.5** Gender equality and inclusion **4.6** Literacy and numeracy **4.7** Education for sustainable development and global citizenship **4.a** Learning environments **4.b** Scholarships } Means of implementation targets **4.c** Teachers

Source: Montoya (2019, Slide 2)

To introduce PISA and PISA-D, the OECD (re)framed the agenda by *narrowing down* the diverse stages to only secondary education and by *limiting* the scope of educational outcomes *to* what is most relevant to PISA; although it has to be noted that with regard to what PISA measures, the OECD is chameleon-like in the use of terms which range

from 'literacy and numeracy', 'universal basic skills', and 'knowledge and skills', to 'learning', 'quality', 'competencies', 'cognitive skills', and '21st century skills', to name a few. The following extracts from the *OECD's Post-2015 Reflection Series* are illustrative of how these operated:

> [1] Experience since 2000 has underlined that schooling doesn't necessarily produce learning. Future education goals should incorporate meaningful and realistic objectives. We have also learnt that in order for education to support social, economic and development outcomes in the 21st century, higher levels of learning will be needed – the kind of competencies and higher order thinking skills that are achieved *only through quality secondary education*.

> [2] OECD is *well placed* to contribute to thinking about future global education goals. In particular, the experience of […] *PISA* that the OECD has implemented since 2000 in partnership with participating countries *is particularly relevant to* a focus on learning in the post-2015 framework. (2013c, 1–2, my emphasis)

In this case, while it makes sense that the post-2015 education agenda should be focusing on learning, I contend that this is the result of deliberate articulation. As Foucault (1972; 1979) and Phillips and Hardy (2002) suggested, discursive practices delineate and constitute what can be seen, said, known and thought. The pragmatic effect is what can be *done* (i.e. the power of discourse). In this case, the vision of SDG 4 (along with its targets) was dealt with in a way to address the *need for measuring* (thereby the need for OECD policy instruments and PISA/PISA-D, in particular). In other words, the agenda was reframed to *prepare for* the authority of OECD expertise, and it served to translate the more complex issues/debates around the definition of SDG 4 (see e.g. King 2017; Unterhalter 2019) into major technical problems that needed to be addressed. As the OECD persisted in suggesting,

> [3] […] a post 2015 education-related goal is likely to include a stronger focus on learning and incorporate the secondary education level. *This kind of goal* will *present* the international community *with* a *major challenge* to *develop* or *identify* and agree on *a universal learning metric*. *How do we define a learning goal that can be measured and tracked over time*? *How do we identify and collect the evidence needed to measure progress*? *What targets* can be set to guide progress towards this goal? Major **OECD policy instruments**, such as the *Programme for International Student Assessment*, have *pioneered new and highly collaborative ways* in which to measure progress in societies *on a global scale*. (ibid., 5, my emphasis)

Here, it is also helpful to cross-refer documents and interviews. The following extracts were taken from two related interviews – Addey's (2017) interview of OECD staff on PISA-D, and my interview conducted on 23 October 2018:

> [1] We positioned ourselves very early on as a voice arguing that the post-2015 world has to be focused on learning outcomes and quality. We cannot have another fifteen years where all talk about this, getting kids into school. And of course, that immediately begged the question 'How

are we going to measure quality, how are we going to measure learning outcomes?'. So, we offered PISA as a global metric, and said 'Look, the work on PISA for Development is going to help us make this available to a wider range of countries'. (OECD2015#30, cited in Addey 2017, 11)

[2] In the time before PISA, the focus was much more on quantity of education: qualifications, degrees. The quality focus was to various aspects that PISA brought into the discussion … There have been measures in the past, but I think to have a comparative collective approach to assessments is quite new development… [and so] now the OECD work places a quite important role at UNESCO, in the measurement of the SDGs. (OECD Interview #2, 2018)

These extracts are illustrative of the OECD's strategy, which operated first to *define* the problem (that SDG 4 should shift its focus to learning and/or quality); it then singled out the lack of a *secondary level*, *internationally comparable*, *learning* metric, and offered PISA and PISA-D as the optimal solutions. What the OECD referred to as 'new and highly collaborative ways' are the adjusted PISA instruments for measuring learning in low-and middle-income countries used in PISA-D (see Chapter 6 for the outcomes).

From the perspective of policy framing, the strategy enacts the 'process of culling a few elements of perceived reality and assembling a narrative that highlights connections among them to promote a particular interpretation' (Entman 2007, 164). In this case, the OECD's (re)framing of SDG 4 renders the other stages of education (such as pre-primary, vocational, higher and adult education) irrelevant; so do the other dimensions of learning that focus on learning environments, scholarship and teachers. I thus agree with Morgan (2017, 19) that 'the framing process is not neutral – it is political as certain issues are foregrounded while others are backgrounded'. Moreover, the strategy also embodies what Biesta (2008) calls the 'normative validity' of measurements – that is privileging what can be measured rather than measuring what we truly value.

It is important to highlight that SDG 4 is composed of seven major targets (from 4.1 to 4.7, as shown in Table 5.2). Ensuring relevant learning outcomes at the lower secondary level is only one of the three indicators for Target 4.1. In 2016, the UIS Global Education Monitoring (GEM) Report team issued a clarification, noting that 'the indicators [that the OECD] used to rank [countries at *Education at Glance*] are not the global indicators but the OECD's interpretation. This nuance is currently missed' (para., 5). The GEM Report team insisted that 'many of the [SDG 4] indicators are not yet sufficiently precise' and hence have been classified into three tiers. According to the classification, the indicator for Target 4.1, the '*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*'

has been classified by the United Nations Statistical Division as a tier III indicator, which means that it has 'no established methodology' (para., 8).

As the Report team further explained,

> [i]n other words, a common standard agreed between countries on what counts as 'minimum proficiency level' is still missing. Indeed, the UIS, as custodian agency, has set up a mechanism for countries to reach consensus on standards for this as well as the other global indicators.
>
> Yet, Education at a Glance has presented results from PISA, its learning assessment of 15-year-olds, *as if it is the accepted standard* for its member states *and beyond to monitor* Target 4.1 – and the same is true of most of the other indicators. It would have been clearer if these had been presented as *potential* options. Or if the proposed approach had been presented explicitly as a regional, *rather than global*, approach to monitoring SDG 4. (ibid., my emphasis)

However, by 2016 (as mentioned earlier), the OECD had already begun to articulate that PISA and PISA-D would serve as the universal monitoring tool for Target 4.1. This is evident in the statement given by Michael Ward, when he was interviewed at the UNESCO IIEP (on 19 February 2016) on the rationales for PISA-D:

> […] now with the SDG in place – Goal 4 for education, there are 10 targets for that education SDG, 7 of them are targets and 3 are means for implementation. You take those 7 targets and 5 of the 7 targets are focusing on learning outcomes and skills, and so the indicators that are developed alongside those targets are emphasising learning outcomes, and the international community is looking for a means for measuring progress against students' pace. I would highlight indicator 4.1, which is levels of proficiency in reading and mathematics, at the end of primary and at the end of lower secondary. And *at the secondary, that's PISA*. And so, we see this PISA for Development and PISA together as a major contribution towards the measurement *on a global scale*, a progress towards that indicator. And when we look at 2018, taking all the countries in PISA, all the countries in PISA for Development, it will be 90 countries taking part in PISA in 2018. And we have enough countries lined up to come into PISA from 2021, it will be something 110, 115 countries in PISA, and so this is […] 90% of the world's population probably covered by PISA at that point. (my emphasis)

This extract is a powerful illustration of the OECD's true ambition for developing PISA-D, which is for PISA to reach universal coverage. It is this ambition that has driven the OECD to actively engage with UNESCO and other related UN bodies so as to define *and* measure the learning targets. To put it differently, the OECD's proclaimed contributions to the post-2015 agenda (identified in the previous strategy) are not ends in themselves but the means to gain the perceived sociopolitical acceptability – with the ultimate outcome of universal participation in PISA. The moral claims associated with ensuring learning for all prepare for the legitimacy of OECD expertise in providing assessment for all.

At the PISA-D International Seminar in 2019, although the OECD declared PISA-D a success, it also admitted that 'when this idea was started actually, it was […] not really

talked' (Schleicher 2019). This implies that the initial phases of PISA-D implementation (as shown in Table 3.2) had not been all that smooth. Indeed, in retrospect, the initiation of PISA was similarly difficult. Schleicher (2018) revealed how he had started by 'piloting' PISA to help clear the path to its wider implementation:

> When I proposed a global test that would allow countries to compare the achievements of their school systems with those of other countries, most said it couldn't be done, shouldn't be done, or wasn't the business of international organisations … I handed my boss [...] a yellow post-it note saying: "Acknowledge that we haven't yet achieved complete consensus on this project, but ask countries if we can try a pilot." The idea of PISA was born. (ibid., 17–18)

In the situation of PISA-D, the OECD had overcome resistance from other influential agencies, as well as from the countries. For example, Auld, Morris and Rappleye (2019) show that UNESCO and UNICEF were among the most vocal critics of Cambodia endorsement of PISA-D, because they saw little value in drawing global comparisons to address the longstanding challenges that the country faces. These accounts reflect the general concerns about the extension of ILSAs into low- and middle-income countries. The OECD again employed the 'piloting' strategy to locate PISA-D within a process of investigative science to avoid criticism (see Auld, Li and Morris 2020). In Chapter 6, by examining the outcomes of PISA-D, I show how, in reality, the experimental success of the PISA-D pilot was preordained.

Moreover, in the field of global educational governance, competition between these key agencies is intensifying (Tikly 2017; Edwards et al. 2018; Elfert 2018). Attempts to persuade, bargain, and otherwise shape the actions of others are likely to encounter varying degrees of resistance or co-optation (Barnett and Finnemore 2004). Organisational legitimacy thus has to be constantly seized. In this case, for the OECD to 'gain' legitimacy for monitoring SDG 4 by overriding resistance and criticism, apart from its historical use of pilots, it would have to dedicate itself to advocating or promoting its authority and expertise, creating a platform for PISA-D to be perceived not only as *reliable* but also as *in demand*. I demonstrate these strategies below.

5.3 Establishing credibility

According to Suchman (1995, 588), organisations seeking to gain legitimacy 'rarely can rely on purely dispositional appeals, because assumptions of good character generally require an established record of consistent performance'. The OECD capitalised on PISA

to this end. There are two aspects relating to this strategy. First, the OECD used a specific set of terms to describe PISA to emphasise its established record of authority and credibility in measuring student learning outcomes. Second, the OECD declared that PISA is *already for development* – that it is already helping to improve education systems and student learning, so that PISA-D is established 'to make the benefits of PISA available to a broader group of countries' (as Schleicher stated at the Initial Technical Meeting in 2013) – and it further selected certain low- and middle-income countries as exemplars to support this claim.

In the case of specific sets of terms, I have identified the following dimensions of terminology deployed by the OECD to describe PISA. In addition to highlighting the *collaborative* efforts that PISA embodies, the OECD also employed words pertaining to the *validity* and *reliability* of PISA instruments and data. Moreover, I have also noticed that the OECD emphasised the *scale* of PISA coverage to further reinforce its prominent position. Concrete examples of these dimensions are evident in the following extracts from the OECD's efforts to promote PISA (before introducing PISA-D, and after having established affiliations and reframed the agenda):

> [1] Shaped by the *shared*, policy-driven interests of participating countries, PISA provides the *most comprehensive* and *rigorous* [*validity*; *reliability*] *international* [*scale*] assessment of learning outcomes (mathematics, reading and science) in education. In keeping with these *common* interests, the scope and nature of the PISA instruments are developed by *leading experts* [*validity*; *reliability*] in the participating countries. (OECD 2013c, 2, my emphasis)

> [2] The OECD is well placed to contribute to this effort, based on its experience with the *highly collaborative* […] PISA. PISA provides *comprehensive* and *rigorous* [*validity*; *reliability*] *international* [*scale*] assessments of learning outcomes […]. Every three years, about *half a million* 15-year-olds from *around 70 countries* [*scale*] are tested under the PISA programme. PISA also collects student, school and system-level *contextual* information, which allows it to *identify factors* associated with quality and equity in learning outcomes [*validity*]. (Davidson, Ward and Palma, OECD 2014, para. 4, my emphasis)

> [3] PISA is *already helping* to measure and improve learning outcomes in *over 70* countries [*scale*]…since its launch…PISA has become *a leading reference* [*validity; reliability*] on the quality of education systems *worldwide* [*scale*]…PISA is *a powerful tool* for policy making. Participating countries receive a *comprehensive* assessment of the quality and equity of their education systems…the *global*, *collaborative* PISA network [*scale*; *credibility*] also helps… (Ward and Zoido, OECD 2015, 21–22, my emphasis)

> [4] Since its launch in 2000, the OECD's […] (PISA) has become a *rigorous* and *comprehensive* [*validity*; *reliability*] *international* [*scale*] assessment of student learning outcomes, serving as *premier* yardstick [*validity*; *reliability*] for evaluating the quality, equity and efficiency of schools. … Students representing *more than 70 countries and economies* that together make up *over 80%* of the world economy [*scale*] have participated in the assessment since its launch, including 44 middle-income countries, 27 of which are recipients of foreign aid. In an effort to make PISA more

accessible and relevant to a wider range of countries with lower performance expectations, the OECD has initiated the […] (PISA-D) project… (OECD 2016b, 3, my emphasis)

When meticulously examined, the terminology ascribed to PISA is highly repetitive and self-reinforcing, creating 'a closed system of circular knowledge' (Silova and Brehm 2015) that bolsters the 'perceived' credibility of the enterprise. By raising the reputation/prestige accorded to PISA, the OECD hopes to establish credibility for PISA-D.

However, considering that the OECD has positioned PISA-D as a *differentiated* (i.e. 'in new and highly collaborative ways') development project targeted at low- and middle-income countries, it would have to further demonstrate how PISA could add value to these new contexts. The second aspect of the strategy is thus critical because the OECD is well aware of the issues and scepticism about PISA for low- and middle-income countries (e.g. PISA's first entry into India in 2009 was also the end of its beginning; see also Pritchell 2012). This awareness is clearly reflected in the initial project document stating that

> […] more needs to be done to encourage greater participation from developing countries. In particular, MICs and LDCs often regard PISA as a survey for rich countries, one that only OECD countries and emerging economies are able to benefit from. There are three particular barriers that need to be overcome if participation in PISA by developing countries is to increase. The first relates to the view among LDCs that their performance in PISA will be so close to the floor in terms of international comparisons that their participation and their PISA results will not yield any policy value. In addition, a poor showing in PISA has political implications for governments and if more developing countries are to participate, ways must be found to help their governments explain to their populations the reasons why their country achieves only a low score in the tests. This will require a good understanding of the political economy of developing countries and the development of methods for managing the presentation of relatively poor results. (2013e, 17)

In this respect, the OECD drew on specific countries, such as Brazil, Peru and Vietnam, as exemplary cases to show how PISA has driven the improvement of learning outcomes in these countries: 'The example of Brazil and other middle-income countries that have successfully managed the presentation of low results on PISA can be drawn upon in this regard' (ibid.).

The experiences of these countries were not only recounted in the OECD documents (as given below), but they were also rehearsed and presented at the PISA-D technical and/or advisory meetings and workshops (see the 6th strategy identified). The OECD had even invited representatives from such countries to attend these onsite events. Among these countries, Brazil was cited most to help support the OECD claim. For example, at the Initial Technical Meeting (2013a), after having listed the benefits of participation in

PISA, Schleicher stated that '[m]any country examples can be cited but Brazil presents an excellent example of how a country has leveraged its participation in PISA to improve learning outcomes and we will be hearing more about this country's experience later today'. Below are some extracts from the PISA-D pamphlets and blogs:

> [1] *Participation in PISA has helped Brazil to improve its education system*. When Brazil first participated in the assessment in 2000, its performance ranked lower than any other country. Brazil used PISA results to benchmark progress, prioritize policies and inform its new national assessment framework. By 2009, Brazil's average PISA scores had improved across the board – and particularly in mathematics, where the *improvement was equivalent to almost one year of schooling* – making it the fastest improver of all PISA-participating countries. (OECD 2013c, 2, my emphasis)

> [2] Twenty-eight low-income and-middle-income countries have participated in PISA. Countries such as *Brazil, Peru and Vietnam have shown how valuable PISA can be* by using the surveys to set quality-of-learning benchmarks and monitoring progress against these over time. As the latest results of PISA 2012 show, the programme has not only helped to identify some of the world's top performing and most equitable education systems; results also show that *diverse* countries have managed to raise the quality of educational outcomes substantially, *despite starting from different points*. (Davidson, Ward and Palma, OECD 2014, para. 5, my emphasis)

> [3] Monitoring and evaluation is crucial for improving education. PISA is a powerful tool because it gives countries an honest assessment of whether their students are on the right track. Brazil has done more than any other PISA participant to improve its education system. Brazil was on the bottom of the ranking when it first participated in 2000. Then Brazil *used PISA* to prioritise policies and to focus more on what policies work better. *Brazil improved the quality of its education system faster than any other nation* over the next 10 years. (Solheim, OECD 2015, para., 7, my emphasis)

By drawing on selected country examples, the OECD employed a particular type of 'persuasive storytelling', which involves 'strategic namings and framings, inserted into a specific context where actors are predisposed to a certain range of policy option' (McCann 2011, 115). In this case, the strategic use of 'namings and framings' had the effect of persuading low- and middle-income countries that if they follow Brazil (by participating in PISA and using the results to inform their national assessment frameworks), they would receive the similar benefits, including improved rankings and higher scores on PISA. However, as discussed in Chapter 2, this logic is problematic in at least two aspects: first, from the CE point of view, it assumes that education systems within countries are all commensurable and, therefore, education reforms and practices implemented elsewhere (say in Brazil) will also work at home; second, from the development perspective, there is a misconceived *methodological equivalence* by which the OECD equates quality of education and learning with improved PISA scores or performance. Consequently, progress is both defined *and* measured by PISA; PISA as a means then becomes an end.

Furthermore, scrutiny of these countries' experiences with PISA suggests a rather more complex picture than the OECD portrayal. For example, according to Silva (2019), Brazil's link to the OECD and its continuous participation in PISA have been strongly influenced by the local agent INEP (the National Institute of Educational Research), and there is a marked difference between what was planned and what was actually achieved. Specifically, she notes that participation in PISA has resulted in major changes 'in the production of information about education more than in the education system itself', and agents did not recognise change at the school and student level (2019, 127). With regard to the PISA situation in Vietnam, one of my interviewees revealed *how*, as a poor country, Vietnam had performed 'really well' on PISA:

> So, for example, Vietnam as an example of how a poor country did really, really well. They have a huge out-of-school population, a huge population! So, we are really just testing their elites. So that means, that one out of two kids don't get to school. And you choose the best student in your family to go to school and the rest work to make it happen. (OECD Interview #1, 2018)

The findings and the statement contradict the OECD portrayal of the successful experiences of such countries engaging with PISA and thus the credibility of PISA and PISA-D. Grek (2013, 695) points out that the production of shared narratives and agendas among experts and policy makers is the result of 'undertakings that often fabricate and manage, rather than strive for a "real" consensus'. In this case, the OECD is not only good at bringing experts and policy makers together around the same table, but it is also efficient at steering and directing them towards its own preordained agenda (see also Chapter 6). It is thus not difficult to imagine that at the PISA-D meetings, country representatives would not give the full picture of what had really happened because that would immediately undermine the OECD claims (that was *not* the purpose of inviting them) and also damage their own image (in front of an international audience). The fact that they would not be opposed, at least openly, to the OECD's statements has allowed it to go further by articulating a demand-led story, as examined below.

5.4 Creating demand

In the previous sections, I showed how the OECD managed to shape post-2015 education discussions and introduced and promoted PISA-D through strategic framings. The OECD started by affiliating itself with the UN agencies specialised in global education. Having gained the perceived sociopolitical acceptability, it then played an active role in shaping

the agenda, arguing that it has to be focused on learning (outcomes), especially at the lower secondary level. In this way, it carved out a niche market for measuring learning, as discussions were soon shifted to the question of how to measure progress on a global scale. It followed that the OECD offered PISA and PISA-D as the optimal solutions to serve as the global learning metric. To leverage its expert authority, the OECD also worked to establish credibility for PISA and PISA-D, by mobilising countries such as Brazil, Peru and Vietnam as persuasive examples of how such countries had benefited from participating in PISA.

Considering the OECD ambition of universal participation in PISA, and in view of UIS' clarification before 2019 of the lack of a common standard for defining and measuring Target 4.1, the provision of PISA and PISA-D is essentially a supply-side story. The OECD is keen to have them adopted by the international community, rather than the other way around. The 4th strategy is about the OECD articulating a *demand-led* story in promoting and legitimating PISA-D. This is evident in the following extracts in which the OECD claimed that PISA-D was introduced *primarily in response to* demand. Two such examples of demand are discernible: one was the demand from low- and middle-income countries who wanted to engage with PISA and enjoy the benefits/prestige accorded; the other was the need of the international community for a universal learning metric to track progress on SDG 4.

To illustrate how this strategy was implemented, it is again helpful to look at some examples. The first two are extracts from statements given by Michael Ward in 2013 as part of the *OECD's Post-2015 Reflection Series*, and in 2016 when he was interviewed at the UNESCO IIEP on the drivers for PISA-D:

> [1] PISA has been one of the most successful education programme at the OECD. We want to *make* the *benefits* of this programme *available to* developing countries…PISA for Development is an initiative that we began *really partly in response to* the post-2015 agenda, which is prioritising obviously education at the centre of development and within education a focus on access plus learning. And of course, *measuring a learning goal is a challenge for the world*. And we consider that *the experience with PISA is very valuable* and if we can increase participation of developing countries in PISA. Then PISA becomes more and more viable as a metric, *a single universal metric* for measuring progress towards the learning goal, post-2015. (2013c, my emphasis)

> [2] The *strong motivation* is *countries really do want to* benchmark their systems internationally, that is a *very powerful driver*. They *want to be* part of the international community around PISA that is focused on learning outcomes and improving learning outcomes. They also *seek to gain access to* assessment experts…PISA is not just about the OECD of course, it is about the countries, and literally hundreds of experts that come together to review frameworks, develop items, produce the tests and then implement PISA. And so, a strong motivation for middle-income countries to come in is to gain access to that.

> [...the first driver is] that more and more many middle-income countries, and low-income countries are coming into PISA, and we are getting a lot of *requests* from developing countries. So, there is a *demand* driver. And the second driver is that the [...] Strategy on Development is taking the best of the OECD policy implements and adapting them to developing countries, making them available to developing countries, the policy implements and the approaches that the OECD countries are using, *for the benefit of* developing countries. And then the third driver is the post-2015 agenda itself... (2016, my emphasis)

These narratives were rehearsed in the blogs and PISA-D technical reports:

> [3] [...] as the number of countries joining PISA *kept rising*, it *became apparent* that the design and implementation models for PISA needed to evolve to successfully *cater to* a larger and more diverse set of countries, including *a growing number* of middle-income and low-income countries *who want to* participate in the assessment. The OECD remains *committed* [...] in maintaining and developing PISA as a global yardstick for measuring success in education. (Schleicher and Costin, OECD 2015, para.3–8, my emphasis)

> [4] Participation in PISA by non-OECD countries *is growing* and is combined *with demand from these countries* for innovations that will maximise their benefits from participation in the assessment. PISA for Development (PISA-D) is an initiative that has been developed *in response to this demand* and *in the context of the Sustainable Development Goal* that was adopted by the United Nations General Assembly in 2015 and which emphasizes universal access to literacy and numeracy. (OECD 2018b, 3, my emphasis)

Similarly, in the *Project Completion Report*, the OECD stated,

> [5] [b]uilding on the experience of middle-income countries in PISA since 2000, and in an effort to *respond to* the *emerging demand for PISA* to *cater for* a wider range of countries, the OECD launched the PISA for Development (PISA-D) initiative in 2014... The project *contributes to the monitoring of international educational targets* related to the Education Sustainable Development Goal (SDG), adopted by the United Nations in 2015 as part of the Agenda for Sustainable Development. (2019b, 7, my emphasis)

However, the OECD documents are lacking nuance in two areas related to demand. Firstly, as the GEM Report team clarified, PISA was considered as one of the *potential* options for monitoring SDG 4; in fact, it was not until 2019 that PISA was officially recognised by UIS as a source of data for measuring progress against, specifically, Target 4.1.1(c). Secondly, the second driver that Ward (2016) referred to as the OECD *Strategy on Development* (examined in Chapter 4) was not provided in the published PISA-D documents; therefore, the effect is that PISA-D was developed primarily in response to the demand from low- and middle-income countries for being part of PISA and from the international community for a global learning metric.

Indeed, these two aspects of demand went hand in hand, because under the OECD narrative, being part of PISA *is* being committed to improving learning; participation is the message for showing commitment and ensuring quality as prescribed in the global education agenda. Accordingly, the OECD not only succeeded in creating a demand-led

story, but also created the self-image as being committed and doing good – by aligning its technical expertise in measuring learning with promoting socially valued goals.

Although one interviewee explained that this is because countries tend to 'vote with their feet', which has made PISA-D as a demand-driven instrument, another interviewee revealed that the OECD approach to these countries has to be *proactive*. I contrast their statements below:

> [1] Because I think the countries themselves, in a way the countries vote with their feet. We *never had intended PISA to become a global instrument*, but once PISA existed, *more and more countries want to be part of that.* For example, we *never invited* countries to take part in PISA, so writing to countries "Would you like to do this?". Countries *always asked* the OECD and sort of the *demand-driven* instrument, and it's a voluntary instrument, financed by the countries, so that is an evolution that happened quite rapidly. (OECD Interview#2, 2018)

> [2] I am pretty sure that *there must have been formal conversations* telling them about the [PISA-D] project. And once we had a basic mind, we *must have approached all the countries* that we thought could be interested in it and of "Would you like to join or not?". And there were several who up to the very last minute, thought "We will participate" and then finally they couldn't. (OECD Interview#1, 2018)

According to Suchman (1995, 587), 'legitimacy building is generally a proactive enterprise, because managers have advance knowledge of their plans and of the need for legitimation'. I find that this proactive dimension resonates with the OECD strategy for articulating that it was the *demand for* PISA that has driven its expansion and, therefore, the introduction of PISA-D (rather than the OECD itself pushing it forward). Moreover, this strategy also embodies what Sellar and Lingard (2014, 722) call 'a self-perpetuating dynamic', which, in this case, is mainly used to help consolidate the status of PISA and mobilise support and legitimacy for PISA-D.

Although there have been contradictory responses (or different stories) from the interviewees, the narratives presented in the OECD publications and blogs are generally consistent and support d'Agnese's (2015, 58) point that PISA is 'more of a life brand than an assessment tool, and one which makes expansive claims'. Following this logic, the next strategy is 'brand' building which treats participation in PISA-D as a sort of membership card that allows for low- and middle-income countries to join the 'global PISA community' to access all the benefits and prestige accorded.

5.5 Membership and belonging

The strategy on membership and belonging echoes Nóvoa and Yariv-Mashal's (2003, 427) notion of 'the politics of mutual accountability', by which they mean a sense of sharing and participation, inviting each country to a perpetual cycle of comparison and improvement. In the case of PISA-D, the OECD emphasises membership and belonging to the *global PISA community*. By being part of this community, countries, especially low- and middle-income ones, can enjoy the benefit (and prestige) of comparing and exchanging with other members of the community (as 'global peers'). Apart from international benchmarking and peer-learning opportunities, low- and middle-income countries can also have access to high-level technical expertise provided by specialists and test experts around the world (but are in fact mainly based in the Global North). Equally important, given that the OECD has positioned the members of the community as a group of countries that value and commit to improving learning and education quality, by being part of the community, low- and middle-income countries demonstrate similar commitment and public accountability.

For example, at the Initial PISA-D Technical Meeting (held in Paris in 2013), Schleicher explained this strategy, summarising the three major benefits of participation: (i) 'being part of…'; (ii) 'good practices and peer learning'; and (iii) 'policy/system reform impacts':

> There are some of the benefits of participating in PISA for developing countries, [i] *being part of* an evidence-based international and national discussion of how to improve student learning outcomes. [ii] Participants in PISA are exposed to *world-class* assessments and rigorous *international* standards. PISA allows countries to learn policy lessons from other countries […], thereby facilitating *peer learning* and the dissemination of *good practices* and knowledge of *what works* to improve student outcomes. [iii] Further benefits of participation relate to the impact of PISA on policy and system reform. [… There are] five areas where we consider PISA has had an impact over the last 10 years… (2013, Slide 9–11, my emphasis)

Similarly, at the PISA-D Capacity Development Workshop (held in Cambodia in 2016), Ward (2016, Slide 2) presented benefits such as 'international benchmarking – Education SDG', 'learning assessment results which better describe the country's levels of proficiency', 'building capacity in assessment', 'join international community focused on learning outcomes with equity', and 'gain access to the world's leading assessment specialists and expertise', as well as 'work with OECD to produce national PISA reports'.

Again, these narratives are repeated in the PISA-D pamphlets. The extracts below are just two examples:

[1] Through PISA results, policy makers can gauge the knowledge and skills of students in their own countries *in comparison with* those in *other* countries, set policy targets against measurable goals achieved in *other* education systems, and *learn from* policies and practices of countries which have *demonstrated improvement*. This kind of *international benchmarking* is more relevant now than ever, given that every country in the world has signed up to the Education Sustainable Development Goal agenda. … PISA-D helps to *build the capacity* of participating countries to conduct large-scale learning assessments and analyse and use the results to support national policies and evidence-based decision making. PISA-D countries also engage in *peer-to-peer learning* with countries *already participating in* PISA. (*PISA-D Brief 2*, OECD 2016e, 1, my emphasis)

[2] An important and *innovative* component of PISA-D is the introduction of *peer-to-peer learning partnerships between PISA-D countries and experienced PISA countries*. In addition to bilateral partnerships […], peer learning has been a feature of PISA-D meetings since 2015. Representatives from Belgium, Brazil, Canada, the Dominican Republic, France, Korea, Kosovo, Luxembourg, Mexico, Spain and the United States have generously volunteered their time and support, *sharing experiences* with PISA-D country representatives… Countries also look to *other* education systems *around the world* to identify *best practices* and *lessons learnt*. (*PISA-D Brief 27*, OECD 2018f, 1–2, my emphasis)

These extracts show that there are two distinctive dimensions added to PISA-D apart from the usual prestige accorded to PISA. One relates to *capacity building* of developing countries to carry out ILSAs and to analyse and use the results, and the other associates *peer-learning* with the *existing* members of the community (most of which are high- and upper middle-income). I suggest that these dimensions are also where the 'banal imperialism' (Silova and Auld 2019) unfolds, considering the implicit assumptions of scientific superiority and statistical 'othering' (explained below).

Firstly, regarding *capacity building*, the OECD's approach to improving countries' capacity is primarily through international contractors who are responsible for survey operations, management and implementation. In these processes, the contractors ensure that PISA-D countries complete 'complex tasks' required of them in order to 'advance' from one phase of the project to the next while 'adhering to PISA's technical standards'; they also 'closely monitor' their progress against 'clearly defined tasks' at each stage of survey implementation (OECD 2017b, 2). In this way, the OECD asserts that '[l]earning-by-doing works: *simply by participating in* PISA-D with its *well-established* and *high-quality* procedure*s* and technical standards, the countries *have acquired valuable knowledge and understanding* of how to manage a large-scale assessment' (ibid., my emphasis).

We understand from the OECD's earlier assertions that PISA-D was supposed to be the outcome of PISA *adapted to suit the needs of* low- and middle-income countries. However, in this case, the statements suggest that PISA had already succeeded in

fulfilling the countries' needs due to its 'well-established' and 'high-quality' procedures and standards. So the implication is that PISA-D countries would benefit simply by adhering to PISA's technical standards, and that there is no point in having additional inputs from these countries because the existing standards are already sufficient and scientifically proven. Indeed, as will be illustrated in Chapter 6, all the undertakings of PISA-D are dominated by the OECD and its contractors. Apart from overseeing the survey design, operation and implementation, the OECD is also tightly in control of the analysis of PISA-D data and the production of national reports. Here, it is helpful to reflect on the implicit assumptions underlying this strategy of membership and belonging, in which the benefit of capacity building is regarded as critical.

Secondly, in the case of *connectivity* and *peer-learning*, I suggest that these extracts illustrate well the OECD's emerging mode of global governance which is being operated through 'circuits of movement and mixture' (Hardt and Negri 2001, 1999). Rather than PISA being a club for exclusive rich (or Western industrialised) countries, the idea of *the global PISA community* provokes a strong feeling of 'mutual accountability', based on sharing and participation, learning and improvement, among different types of countries at different stages of development. Therefore, representatives from Belgium, Brazil, Canada, the Dominican Republic, France, Korea, Kosovo, etc. sit with PISA-D country representatives – all mixing together. The important point here is that the OECD's modulation and control are duly legitimised through such processes of *exchange*, *dialogue*, *membership* and *belonging*. There is no danger of identifying 'what works' and adopting best practices, because the choices now are 'diverse' enough to accommodate (or cater to) a wider range of countries – like a smorgasbord with a large variety of enticing dishes from which every country can take its pick.

However, will low- and middle-income countries be better off as part of the global PISA community? It is helpful to first cross-refer to Grek's (2013, 706) interview of members of the European Committee (EC) on reaching 'real' consensus:

> [...] these [supra-national] institutions will benefit from that – the more they [i.e. countries] contradict each other, the more the institutions decide…. And with OECD, surely it is the same. This is so obvious – that's what they do – OECD is Andreas. We always have a joke with Andreas – where he is brilliant, is to conclude. He is fantastic in this – conclusions! He is the conclusions expert – they are in before the meeting (laughs) … It is very convenient.

To be more serious, as Nóvoa and Yariv-Mashal (2003, 428) point out, '[t]he construction of comparable indicators serves as a "reference point" that will eventually lead the various

national institutions to adopt "freely" the same kind of actions and perspectives within the educational field'. In this case, the field of PISA is pre-defined by the OECD with the support of its technical partners (i.e. specialists and test experts). For low- and middle-income countries, there emerges the greater cross-Atlantic knowledge divide (Kaess 2018) as the OECD aims for a universally applicable learning metric that undoubtedly fails to incorporate local and indigenous knowledge into the test (see also Chapter 6). The critical focus here is the *greater inclusiveness* that this global PISA community embodies, and in this way, the OECD invites all to enter 'peacefully' into its sphere of operation. Drawing on Hardt and Negri's (2001, 189) description,

> [t]he Empire does not fortify its boundaries to push others away, but rather pulls them within its pacific order, like a powerful vortex. With boundaries and differences suppressed or set aside, the Empire is a kind of smooth space across which subjectivities glide without substantial resistance or conflict.

I contend that this greater inclusiveness is the magnanimous face of the global PISA community, which sets aside differences (that are inflexible or unmanageable) but maintains a rhetorical acknowledgement of countries' differences and preferences. 'The OECD and its contractors provided countries with *tailored* support to reinforce their capacity and create the conditions in each country that are needed to *successfully implement* the assessment' (OECD 2019c, 38, my emphasis). Accordingly, and not surprisingly, through PISA-D, the OECD declared in the *Project Completion Report* that

> PISA [now] is the *world's most **inclusive** metric* on student outcomes and one of *the most **global** instruments* of the OECD already, with 78 OECD Members and non-Members involved in the 2018 round. PISA is particularly valuable *not just for monitoring* SDG progress but also for *informing policy and funding decisions* that will help countries **achieve** the SDGs. (ibid., 90, my emphasis)

Morris (2015a) anticipated when PISA-D was introduced that there was the prospect of aid being tied to performance in PISA-type tests. We now see that this is partly realised in the ways that the OECD is harnessing SDG 4 monitoring efforts to prevail on countries to allocate more resources and to continue to participate in PISA:

> UNESCO, UIS, GEM Report team and the OECD conclude that to meet the shortfall on SDG 4 monitoring in LMIC of USD$132 million per year, it will be necessary to increase the domestic and external resources allocated to SDG monitoring, especially *to support learning assessments such as PISA* […]. There is no doubt that if this financing gap is not closed, there will continue to be funding *constraints* to low-income country participation in PISA. (OECD 2019c, 101, my emphasis)

In retrospect, the OECD's intention to sustain PISA participation was revealed when Ward spoke to the UNESCO IIEP staff member during the interview conducted on 19 February 2016. I provide his statements below:

> The key thing for us is *sustainability*, and we want countries not just to come into PISA and go out. We want them to stay for a longer […] and to be able to track their progress over time. And so that requires a line on the recurrent budget assessment, participation in international assessment whatever, and that I think is a challenge in certain countries to be able to make structural changes. (my emphasis)

Here, the cloak of SDG 4 monitoring prepares for a universal acceptance of PISA. The moral sense of authority helps to accelerate the OECD's proposal of 'data revolution' in developing countries and, therefore, the infusion of its 'DNA' into the field of education and international development (see Chapter 4).

Last, the OECD discourses, examined above, have been well-rehearsed at the global PISA-D meetings and workshops held along the way. I interpret their roles and movements below.

5.6 Persuasion in motion

Grek (2013) highlights the significance of the travelling expert in constructing a space of equivalence, where policy objectives are shared, and improvement of performance is achieved through constant comparisons. The OECD is a key player in mobilising education specialists, test experts, and policy makers around the world. In the case of PISA-D, the OECD held meetings, workshops, and an international seminar to disseminate and legitimate PISA-D at a global scale. These meetings, workshops, and seminar were 'on-site events' and 'co-present encounters' (Urry 2003, 155) whereby the OECD discourses continued to circulate and gained momentum, and PISA-D deals were done.

Through these events and co-presence, the OECD not only provided opportunities for PISA-D country representatives to network with global peers and test experts (and vice versa), but it also enhanced the reputation of PISA and PISA-D among, and encouraged their adoption by, a wider community of policy actors. Therefore, I consider these events (i.e. the meetings, workshops and the seminar) as the *sites of persuasion* and *persuasion in motion* – moving from Paris to Washington D.C. to Montreal…to Phnom Penh to Asuncion…to Saly and to London (see Table 5.3 for the full range of movements).

Table 5.3 PISA-D legitimation strategy – *persuasion in motion.*

Sites of Persuasion	Movements
PISA-D International Seminar	*London, U.K. – 25th Sept. 2019*
9th International/National Project Managers Meeting	*Antigua, Guatemala – (10–15 day) July 2018*
8th International/National Project Managers Meeting	*Saly, Senegal – 2nd–4th May 2018*
7th International/National Project Managers Meeting	*Princeton, New Jersey, U.S. – 4th–8th July 2017*
6th International/National Project Managers Meeting	*Siem Reap, Cambodia – 22nd– 26th May 2017*
4th Meeting of the International Advisory Group	*Siem Reap, Cambodia – 17th–19th May 2017*
5th International/National Project Managers Meeting	*Madrid, Spain – 1st– 3rd Nov. 2016*
4th International/National Project Managers Meeting	*Livingston, Zambia – 4th– 8th July 2016*
2016 Annual Meeting – PISA-D Presentation	*Washington D.C., U.S. – 12th April 2016*
3rd International/National Project Managers Meeting	*Asuncion, Paraguay – 4th– 8th April 2016*
3rd Meeting of the International Advisory Group	*Asuncion, Paraguay – 30th March–1st April 2016*
Capacity Development Workshop: PISA-D	*Phnom Penh, Cambodia – 29th Feb. 2016*
2nd International/National Project Managers Meeting	*Rockville, United States – 25th–29th Jan. 2016*
1st International/National Project Managers Meeting	*Quito, Ecuador – 28th Sept.–2nd Oct. 2015*
PISA-D Questionnaire Expert Group	*Stellenbosch, South Africa – 20th– 24th Aug. 2015*
PISA-D Subject Matter Expert Group	*Mexico City, Mexico – 20th–24th July 2015*
2nd Meeting of the International Advisory Group	*Paris, France – 11th–13th March 2015*
Technical Workshop on Out-of-School 15-Year-Olds	*Montreal, Canada – 1st–2nd Oct. 2014*
1st Meeting of the International Advisory Group	*Paris, France – 27th–28th May 2014*
Technical Workshop on Cognitive Instruments	*Washington D.C., U.S. – 8th–11th April 2014*
Initial Technical Meeting	*Paris, France – 27th–28th June 2013*

Sources: PISA-D meetings and workshops (2013-2018); PISA-D International Seminar (2019)

As the OECD's 'outreach activities' (2019b, 107), these events can be seen as operating a specific form of marketing strategy which I call PISA 'boosterism' (adapted from 'policy boosterism' in McCann 2011). The strategy of *persuasion in motion* thus refers to the migration of PISA-D 'boosterism' at a global scale. Specifically, as the table shows, *for* each site of persuasion, a specific theme (such as technical, annual, advisory or managerial) is set out, with some key attendants (including e.g. the NPMs, country representatives from peer countries, and international contractors) remaining at a permanent base whereas the others are a fleeting presence (depending on the location of each site and the availability of development partners and local agents). *Before* or *after*, PISA-D pamphlets, technical reports, and working papers are published, blogs are updated, and mission trips and support visits are made, along with a warm feeling of helping the poor (or supporting their 'capacity building', as documented in the publications); and *at* the site, documents are circulated, PowerPoint presentations are given, models explained, experiences recounted, and applause given.

These microspaces are 'pre-eminently social settings and events of speaking and exchange where "trust" is built, commitments are made, and deals done' (Ball 2012, 68). These microspaces are the key relational sites where the sense of membership and belonging to the global PISA community articulated by the OECD *materialises*. As Urry (2004, 28) puts it, 'all social life [involves] various kinds of connections sustained at a distance but with intermittent [face-to-face conversations and] meetings'. In this case, in addition to maintaining connections, these meetings and workshops also ensure substantial progress is made towards the completion and 'success' of the pilot project by establishing 'shared' understanding and 'collaborative' relations between PISA-D countries and experienced PISA participants, as well as international contractors. Peer learning is suggested as 'a feature of every international training event with experienced PISA participants sharing their lessons and experiences with the PISA-D country representatives' (OECD 2017b, 2).

As noted earlier, the purpose of each technical event is to ensure that PISA-D countries complete the complex tasks required of them and adhere to PISA's technical standards. Through these sites of co-presence and mutual learning then, the OECD and PISA-D contractors *closely monitor* countries' progress in predefined tasks so as to ensure consistency and comparability. These ways of operation through training and meetings are illustrative of how the OECD manages differences (between members, especially the

experienced and the new) and hierarchises the community (i.e. placing PISA-D countries, behind in place and time, by means of technical procedures and standards).

Therefore, the events for sharing and learning, co-presence and hands-on training reveal the OECD's continuing hold on imperialist assumptions of dominance, superiority and 'othering'. The discourses that the OECD employ, which I have unpacked through analysing these strategies, have been designed not only to regulate the different types of countries but also, ultimately, to reconstruct them. Within the global PISA community, PISA-D countries are construed as backward and as needing to learn (from experienced members and technical experts) to improve themselves. As Billig (1995, 37) notes in the context of 'banal nationalism', 'ideology operates to make people forget that their world has been historically constructed'. This also applies to PISA-D which it is *not* free from ideology. Through PISA-D, these countries turn into 'a site for measurement' (Cowen 2018), the OECD expropriates their culture and knowledge, 'denying their legitimacy' and imposing its 'own univocal logic' (d'Agnese 2015, 56). The development trajectory is preordained, and PISA-D countries embark on it by climbing the PISA ladder.

Summary

In this chapter, I have interpreted the six major strategies the OECD employed to promote and legitimate PISA-D: (1) *building affiliations;* (2) *shaping and (re)framing SDG 4;* (3) *establishing credibility;* (4) *creating demand;* (5) *membership and belonging;* and (6) *persuasion in motion*. Together, they illustrate the constitutive effect of the OECD discourses. The ultimate effect has been to achieve *assessment for all*, which is to globalise OECD data provision and analysis through engaging with SDG 4 and incorporating low- and middle-income countries into the global PISA community.

Within these discursive practices, change is incremental, with such notions as PISA for 'development', data for 'development' and assessment for all, 'which seemed radical, even unthinkable, become more and more possible, then normal and then necessary' (Ball 2012, 113). This effect is attributed not to the minor undertakings of a single actor, but to the repertoire of strategies and combination of discourses that 'overlap, repeat, or imitate one another according to their domain of application, […] converge and gradually produce the blueprint of a general method' (Foucault 1979, 138). Learning now *has* to be measured for improvement, and PISA now *has* to be there for development. Concealed

within these practices is the 'banal imperialism' (Silova and Auld 2019) of placing low- and middle-income countries remote in space and behind in time. That is to say, they are construed as needing to be put on the PISA ladder and to learn from other members of the community to improve themselves.

The last two strategies of *membership and belonging* and *persuasion in motion* illustrate well the emerging OECD governance approach operating through mechanisms of configuration and modulation (explained in the previous chapter). In particular, the idea of a global PISA community provokes a strong sense of sharing and participation, inclusion and flexibility, inviting all to come peacefully within its realm of control. The microspaces of co-presence and events of speaking and exchange can be read as a microcosm of processes in which the OECD (as moderator) carefully manages differences and facilitates dialogues, and also sustains order, albeit implicitly, within members of the community. Consequently, while the boundary of the empire keeps expanding by becoming more inclusive and diverse, the central positioning of key actors, such as the OECD and test experts, and the underpinning imperialist assumptions and practices consolidate and penetrate. To use Hardt and Negri's (2001, 200) terms, 'divide and conquer' is not really the current formulation of imperial strategy. 'The triple imperative of the Empire is to incorporate, differentiate, [and] manage'.

These three distinctive moments seem to resonate with the general apparatus of the OECD imperial command. Through a critical interpretation of the major strategies, this chapter questions and problematises the taken-for-granted assumptions, ideologies and practices, and thereby the sort of moral legitimacy gained through PISA for development. In the following chapter, I examine the outcomes of PISA-D, focusing on the results generated from the test, the ways in which the OECD compares them, as well as the policy actions it suggests to help improve the education systems of the piloting countries. The third line of inquiry is vitally important given that the OECD has positioned these outcomes as its major contributions to supporting SDG 4, and which have provided the compelling logic for extending PISA into low- and middle-income countries (as seen from the 1st and 2nd strategies). An examination of these outcomes will also help to understand the actual use of OECD's expert knowledge for the enhancement of PISA instruments, the analysis of data and results, and the production of reports which relate to the 'credibility' it has established through the 3rd strategy and the 'benefits' it has proclaimed for developing countries.

Chapter 6. PISA-D Outcomes

Introduction

> PISA for Development (PISA-D), the latest of the OECD's PISA assessments, focuses on making PISA more accessible and relevant to low-to-middle-income countries. We have enhanced our PISA instruments so that they target the range of student performance in these countries. We have also collected background information to capture how students learn, teachers teach and schools operate in these contexts. PISA-D has also helped the participating countries to build their capacity to manage large-scale assessments and to make use of the results in support of national policy dialogue and education policy-making. (OECD 2018c, 2)

In this chapter, I examine the major outcomes of PISA-D, including the 'enhanced' assessment instruments, the findings from the school-based assessments, and the policy recommendations presented in the main OECD report and specific country reports. I also look at the PISA-D International Seminar (2019) which marked the closure of this project and where the major outcomes were announced. These sources are read intertextually with reference to the claims made by the OECD of the challenges that would be overcome and goals that would be delivered by PISA-D. The chapter is significant in two ways. First, the major outcomes serve to compose a larger PISA programme which will be used for all countries that participate in future PISA cycles. The examination of these outcomes will help to understand their future use and the OECD approach to global educational governance post-2015 (see Chapter 4). Second, the outcomes are essential for the OECD's expert authority, which has helped to legitimate PISA-D as the universal monitoring tool for SDG 4 (see Chapter 5). As the OECD (2019c, 80) states, 'through its enhancement of PISA, the PISA-D initiative is designed to inform and support the monitoring, reporting and achievement of the education SDG'.

It is also important to highlight that when PISA-D was first introduced, it was positioned as a 'pilot' project (OECD 2013e, 2014b). The stated objective was to adapt and enhance the existing instruments for low- and middle-income countries while maintaining their comparability with the main PISA standard. At the Initial Technical Meeting (2013), the OECD staff acknowledged that this task would not be easy, which was why they were regarding this as a 'pilot'. Although the process of piloting may help to generate evidence of feasibility and effectiveness, Ettelt, Mays and Allen (2015) contend that policy pilots primarily serve as an instrument designed to achieve the objectives of policymakers, rather than to question or evaluate them. In the case of PISA-D, considering the significance of challenges and the OECD's historical use of pilots to

underline the scientific mode of operation and to overcome opposition to its new initiatives, Auld, Li and Morris (2020) suggest that the experimental success of the PISA-D pilot is preordained.

I concur and suggest that this can be further observed from the long delay of the output of the out-of-school surveys and the publication of the *Project Completion Report* (2019c) without the output from these surveys. For example, in the report, the OECD claimed that 'while this output of the project is not scheduled for achievement until March 2020, it is already possible to conclude on the basis of the work completed to date, that this [out-of-school] component will be successful' (ibid., 65). Recognising the OECD's strategic use of pilots and interrogation of the major outcomes, this chapter sheds light on how PISA-D has failed to resolve the challenges that participating countries confront and on how the stated experimental success undermines its ability to fulfil those tasks.

The chapter is structured as follows: it first focuses on the assessment instruments based on the three technical strands (i.e. cognitive tests, contextual questionnaires, and the out-of-school component), highlighting both the achieved outputs and unresolved problems; it then looks at the findings from the school-based assessments (presented in the main report) and how these are compared and interpreted by the OECD; finally, it examines the 'effective interventions' suggested in the main report and the different types of 'policy actions' promoted in the specific country reports. To better understand the content of the country reports, a meticulous examination of the ways in which these reports were produced and framed is provided *à* priori.

PISA-D assessment instruments

> PISA-D was successful in making the assessment instruments more relevant to low-to-middle-income countries while still being able to report results on the main PISA scale, thus facilitating international comparisons on all of the variables covered by PISA. These variables and factors include student performance, educational attainment, health and well-being, attitudes towards school and learning, the learning environment, learning time, the quality of instruction, family and community support, and resources devoted to education. (OECD 2019a, 1)

The lower rungs on the PISA ladder

PISA-D set out to make the assessment instruments more accessible and relevant to low- and middle-income countries engaging in PISA.

A particular focus was to introduce additional rungs at the bottom of the PISA ladder from which such nations could begin their ascent. In some of these countries, students' prior performance on PISA was concentrated at the lower levels of the test's proficiency scales (i.e. six scales from Level 1 to Level 6). The OECD staff (Avvisati 2019, Slide 4) considered that the test was 'too hard for most students in low- and middle-income countries'. To address this, PISA-D incorporated more test items at the lower end of the scale and created lower proficiency levels, such as Level 1a, 1b, 1c, 2a, 2b and 2c. Out of the total number of test items, these lower proficiency levels account for about 65% while Level 3 and above are for higher proficiency levels. The OECD (2019c) stated that this kind of arrangement helps to achieve two major outputs. First, those 'easier' items can better measure the proficiencies of students in low- and middle-income countries. Second, with a high concentration on Level 2 and below, the findings can inform countries' progress on SDG 4 Target 4.1.1(c), which requires all children and young people at the end of lower secondary to attain the minimum level of proficiency (i.e. Level 2).

Another focus of the PISA-D test item design was to develop an unbroken PISA development scale. The purpose is to ensure comparability between findings generated from the 'easier' items and those from the main PISA assessments. To achieve this, items added at the lower end comprised a large proportion (i.e. 60–70%) of trend items of PISA 2015, which had been accumulated since PISA 2000. The rest of the 'easier' items were drawn from some other assessments, including PISA for Schools, PIAAC (Programme for the International Assessment of Adult Competencies), STEP (Skills Toward Employability and Prosperity) and LAMP (Literacy Assessment and Monitoring Programme of adults' reading and numeracy skills). A schematic overview of the test item composition is reproduced in Figure 6.1.

Figure 6.1 Linking PISA-D test items to the main PISA item pool.
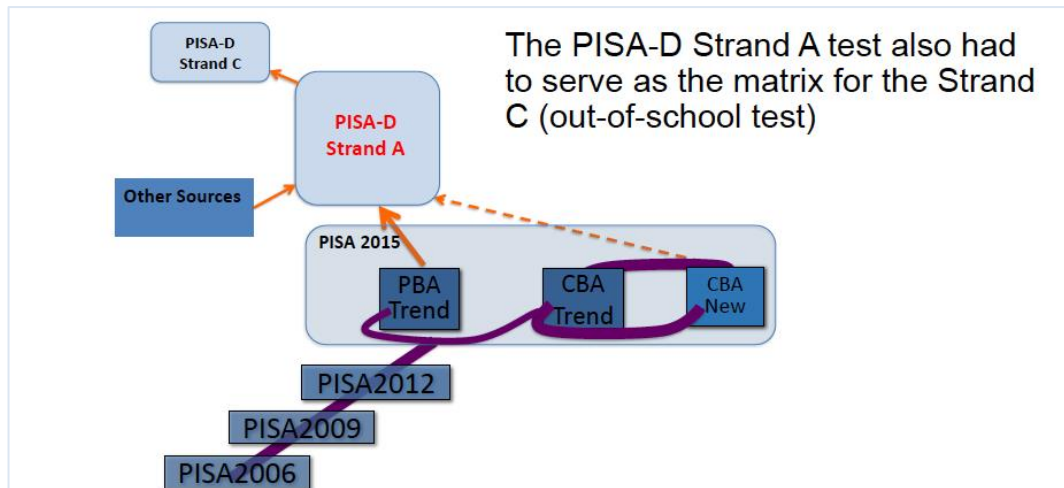
PBA: Paper-based assessment; CBA: Computer-based assessment

Given that the primary purpose was to obtain an accurate link to the PISA scales, it remains unclear to what extent test items from other sources (such as STEP and LAMP) were used. While PISA-D test experts provided the configuration for each domain, as summarised in Table 6.1, the items clustered in PISA 2015 and other related OECD studies. In the main OECD report on PISA-D results, the findings were compared, 'wherever possible' to the OECD average for the PISA 2015 assessment cycle (2018c, 6).

Table 6.1 Distribution of PISA-D test items by source and level.

	Reading	Mathematics	Science
PISA 2015 trend	79%	62%	68%
Other sources (PISA for schools, PIAAC, STEP, LAMP)	21%	38%	32%
Level 2 and below	65%	65%	45%

The arrangements of the cognitive items examined above suggest that PISA-D served to achieve two goals. The first was to prepare for the international comparability of PISA results by establishing a strong linkage of the PISA-D scales to the main PISA (e.g. PISA 2015) and, also, to the SDG 4 indicator of minimum proficiency at the secondary level.

As confirmed in the *Project Completion Report* (2019c, 39), 'PISA-D was successful in making the assessment instruments more relevant to LMIC while still being able to report results on the main PISA scale, thus facilitating international comparisons on all of the variables covered by PISA'. The second was to prepare for universal participation in PISA by constructing one continuous PISA development scale (or trajectory) in which low- and middle-income countries could begin their ascent. In this way, 'this broader set of information about performance at and below the baseline level of proficiency will be directly relevant to curriculum planners, teacher educators and other education professionals in LMIC in designing improvement plans and policies' (ibid., 26). Not surprisingly, in order to ensure PISA-D a success, 'the OECD and its contractors provided countries with tailored support to reinforce their capacity and create the conditions in each country that are needed to successfully implement the assessment' (ibid., 38).

The educational prosperity promised by Willms

In terms of PISA-D contextual questionnaires, OECD adopted the Educational Prosperity Approach (hereafter the EPA) to help enhance their relevance to the contexts of low- and middle-income countries. The EPA was also applied to frame the findings presented in the main OECD report and specific country reports (which I examine below). Specifically, the EPA served to identify additional indicators that are related to students' performance but are beyond school and classroom practices, such as indicators relating to the social and economic context. The latter aspects are captured in the main PISA contextual questionnaires. The additional indicators include, for example, questions about educational attainment, health and well-being and attitudes towards school and learning.

Overall, the PISA-D contextual questionnaires were structured according to the five core 'foundations for success' outlined by the EPA: inclusive environments; quality instruction; learning time; material resources; and family and community support. These 'foundations for success' are positioned as the key factors that will enable improvements across a set of four key outcomes called 'prosperity outcomes': educational attainment, academic performance, health and well-being and attitudes towards school and learning. Figure 6.2 shows how the EPA was imagined as 'The Prosperity Tree' in the main OECD report. This image was also selected as the cover page of each individual country report.

Figure 6.2 The EPA imagined as 'The Prosperity Tree'.
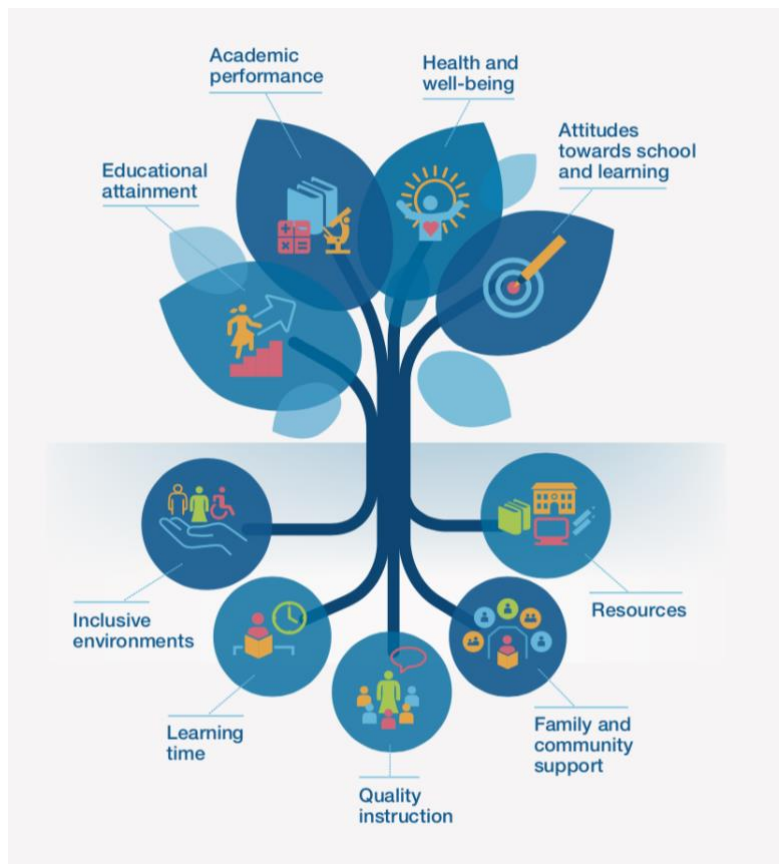
The application of the EPA to PISA-D has largely been overlooked in the critical literature. This approach was developed by J. Douglas Willms, a Canadian scholar based at the University of New Brunswick. He is the founder and president of The Learning Bar – a consultancy that has worked for UNESCO and OECD on aspects of their assessment programmes for many years (see Willms 2003, 2006). Indeed, Willms explained, in an interview, that he embarked on the current career through engaging with OECD PISA:

> In fact, it started more than ten years ago when I was working with a small group that developed the Organization for Economic Cooperation and Development (OECD) PISA study. At that time, I wanted to get some non-cognitive outcomes into PISA. We know that learning is a function of the quality of instruction, but it also requires emotional and intellectual engagement on the part of students.

> It is difficult to convince organizations like the OECD to move beyond the assessment of cognitive outcomes, but we were successful in getting a few measures of student engagement included in PISA – students' sense of belonging and truancy. The OECD later commissioned me to conduct a study that compared levels of student engagement across countries. (*In Conversation*, 2011, 2)

As can be understood from these statements, Willms specialises in child literacy. His organisation provides services in four primary areas: (1) early childhood evaluation and intervention systems; (2) step-by-step skills-based literacy development; (3) school climate and improvement surveys; and (4) professional learning programmes[9]. The overarching approach to address these areas is the EPA. Willms has positioned it as a unique assessment framework which focuses on factors that impact on all aspects of a child's well-being and learning outcomes.

The explanatory logic of the EPA, however, differs from the school effects model that is central to the main PISA, especially the PISA contextual questionnaires. The EPA emphasises the longitudinal, cumulative effects of learning, whereas the school effects model focuses on the cross-sectional, student, family, classroom and school factors. Indeed, Willms (2018b) proposed the EPA as an alternative to the school effects model, criticising that the latter is premised on an input–process–output analytical framework, and, so, isolates the effects attributable to key resources and processes that are outside school settings and accumulated from an earlier stage of children's development. As he elaborated,

> [t]his 'quest for school effects' has been a central feature of all large-scale international studies, including PISA and the IEA studies. The contextual frameworks have been based on an input-process-output 'production function' paradigm, attempting to capture the most salient student, family, classroom and school factors that explain student achievement. The analyses of data have been based on multilevel regression models that examine the relationships between a student outcome, such as reading performance and a long list of school and classroom factors … The main issue is that students' performance […] on the PISA tests at age 15, are the *cumulative result* of countless factors that affect children's development, beginning at conception and continuing through to the time of the assessment. (2018b, 7, original emphasis)
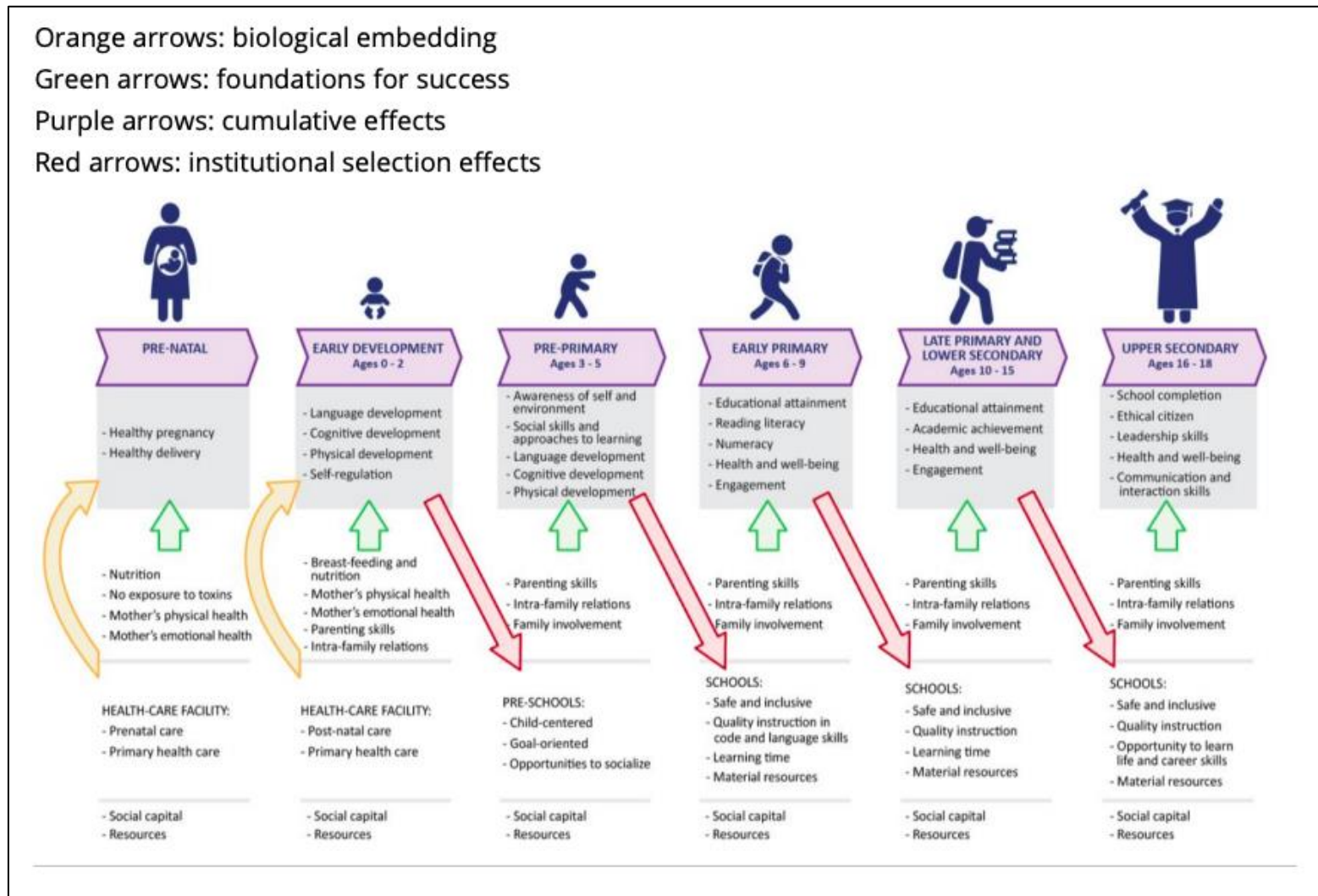
By contrast, the EPA is portrayed to adopt a life-course approach to assess children's learning from the early years to adolescence, with attention to the transitions and enabling resources. Thus, in addition to the 'foundations for success' captured in the PISA-D contextual questionnaires, the EPA also covers 'biological embedding', 'cumulative effects' and 'institutional selection effects', as shown in Figure 6.3.

---

[9] The Learning Bar website, accessed 11 May 2019, https://thelearningbar.com/

Figure 6.3 The explanatory logic of the EPA: four ways that success accumulates.

Orange arrows: biological embedding
Green arrows: foundations for success
Purple arrows: cumulative effects
Red arrows: institutional selection effects

PRE-NATAL

- Healthy pregnancy
- Healthy delivery

- Nutrition
- No exposure to toxins
- Mother's physical health
- Mother's emotional health

HEALTH-CARE FACILITY:
- Prenatal care
- Primary health care

- Social capital
- Resources

EARLY DEVELOPMENT
Ages 0 - 2

- Language development
- Cognitive development
- Physical development
- Self-regulation

- Breast-feeding and nutrition
- Mother's physical health
- Mother's emotional health
- Parenting skills
- Intra-family relations

HEALTH-CARE FACILITY:
- Post-natal care
- Primary health care

- Social capital
- Resources

PRE-PRIMARY
Ages 3 - 5

- Awareness of self and environment
- Social skills and approaches to learning
- Language development
- Cognitive development
- Physical development

- Parenting skills
- Intra-family relations
- Family involvement

PRE-SCHOOLS:
- Child-centered
- Goal-oriented
- Opportunities to socialize

- Social capital
- Resources

EARLY PRIMARY
Ages 6 - 9

- Educational attainment
- Reading literacy
- Numeracy
- Health and well-being
- Engagement

- Parenting skills
- Intra-family relations
- Family involvement

SCHOOLS:
- Safe and inclusive
- Quality instruction in code and language skills
- Learning time
- Material resources

- Social capital
- Resources

LATE PRIMARY AND LOWER SECONDARY
Ages 10 - 15

- Educational attainment
- Academic achievement
- Health and well-being
- Engagement

- Parenting skills
- Intra-family relations
- Family involvement

SCHOOLS:
- Safe and inclusive
- Quality instruction
- Learning time
- Material resources

- Social capital
- Resources

UPPER SECONDARY
Ages 16 - 18

- School completion
- Ethical citizen
- Leadership skills
- Health and well-being
- Communication and interaction skills

- Parenting skills
- Intra-family relations
- Family involvement

SCHOOLS:
- Safe and inclusive
- Quality instruction
- Opportunity to learn life and career skills
- Material resources

- Social capital
- Resources

The four ways elaborated in the framework suggest that PISA-D only selectively adopted the EPA, which constitute the 'foundations for success' (green arrows) at the fifth stage of children's development, and then repackaged it as 'the Prosperity Tree' (Figure 6.2). Surprisingly, Willms, who was opposed to the premises that underpin ILSAs such as PISA, encouraged low- and middle-income countries' participation in PISA-D by suggesting that they 'would benefit by understanding how well their students fare compared with students in other countries' (2018b, 9). However, given that the EPA stresses learning improvements over the whole life-course, it advocates a continuous measuring of 'prosperity outcomes' from the child's pre-school period to the three schooling periods. It thus requires early collections of data in order to establish standards and inform progress for the subsequent stages. This would suggest a far greater emphasis on early childhood, rather than on 15-year-olds.

The application of the EPA to PISA-D contextual questionnaires also raises several critical issues. Here, I reflect on three.

First, as the literature shows, scientific studies on child development largely draw samples from the Global North, which do not represent the indigenous cultural and epistemological diversity in low- and middle-income countries (see e.g. Arnett 2008; Heinrich, Heine and Norenzayan 2010; Wagner, Wolf and Boruch 2018). This especially applies to the EPA, which has its roots in pupil literacy in Canada and Australia. For example, The Learning Bar's school literacy programme – Confident Learners – was developed based on curricula from Canadian provinces and some US states, as well as from Australia and the UK (Willms 2018a). It is thus questionable that the selective use of the EPA would be sufficient for explaining the variance across and within PISA-D participating countries.

In fact, Willms (2018b) assumes the universal reality of his model, arguing that it represents the key markers of child development and the necessary conditions for success at each stage in all countries alike. In this regard, the EPA is an instrument of evaluation 'inscribed with the values not only of universality, objectivity and stability, but also certainty and [thereby] closure' (Moss 2016, 9–10). As we shall see, the EPA soon became the OECD's technical device of 'applying expert-derived templates to particular settings' (ibid.) in the PISA-D countries. Therefore, the outcomes of the contextual questionnaires informed by the EPA merely legitimate the dominant forms of scientific

knowledge, which continue to marginalise local and indigenous knowledge of low and middle-income countries (Kaess 2018).

Second, the EPA is also unsuited for the settings in PISA-D countries in that it requires resources and high-level institutional capacity for its implementation. For instance, Willms (2018b, 10) suggests that the development of children's reading skills should be the primary focus of educational monitoring, as 'it is a pre-requisite for the development of strong academic skills at the lower and upper secondary levels and is essential to school completion and social justice'. This proposal appears noteworthy, but it is unclear how this could be adapted to more disadvantaged families, where their parents may not be able to read (or in which other languages are spoken at home) and where youth are out of school (examined further below). Moreover, the EPA would also require complex implementation strategies and proprietary software provided by The Learning Bar. Lockheed reflects on the practical challenges:

> Ministers of education in low-income countries are aware that children are not learning … Ministers also understand many of the requisites for learning. But their main challenges are ensuring that resources are aligned with desired outcomes and implementing policies related to the financing, deployment, and monitoring of resources. To utilize these assessment tools in the poorest countries, where the need is greatest, may exceed both the budgetary and implementation capacity of these countries. (2018, 118)

Lockheed's point indicates the third critical issue concerning the life-course approach underlying the EPA and its extensive use of data. The EPA emphasises the continuous monitoring of children's development from early childhood to adolescence. The implication is that it has the potential to be fully integrated into the OECD's wider assessment agenda, linking its International Early Learning and Child Well-being Study (IELS) to the main PISA and, even further, up to assessments of higher education learning outcomes and adult skills (e.g. PIAAC).

In fact, the OECD has already acknowledged the impact that children's early learning has on later life outcomes and claimed that IELS would contribute to providing information on the trajectory between early learning outcomes and those at age 15 and beyond (OECD 2015, 2018d). Moss and Urban (2019, 2020) suggest that IELS is an intended part of the OECD's growing global web of measurement, which seeks to give the organisation a pre-eminent position as education's global governor. Considering that the current IELS design still lacks a clear paradigmatic positioning and, as such, has been widely criticised, it is likely that the EPA would serve as a provisional solution.

Third, as children's learning is the result of the cumulative and interrelated effects of 'biological embedding', 'foundational factors' and 'institutional selection', it remains unclear how exactly the 'foundations for success' actually impact on the 'prosperity outcomes' (Figure 6.2). While Willms was critical about the framework used in PISA, many of the examples he drew on to elaborate the EPA are based on PISA data (e.g. the PISA 2015 data for Mexico). As such, it is not evident that the relations between those factors and outcomes were adequately addressed. The EPA's logic seems to suggest that countries which devote more resources to education, provide more inclusive environments, better quality instruction, and greater community and family support will enjoy better educational outcomes. These are self-evident truisms which derive from the meanings of the words employed. Poor quality teaching and excluding pupils by definition will not improve students' overall performance (see also Auld, Li and Morris 2020).

The missing millions: how much can PISA-D help?

Large-scale assessment programmes carried out in low-resourced settings are generally considered inappropriate for obtaining information about the poorest. This is because 'they omit *by design*: the homeless; those in institutions [such as refugee camps]; and mobile, nomadic, or pastoralist populations' and '*in practice*, typically under-represent: those in fragile, disjointed households; slum populations and areas posing security risks' (Carr-Hill 2013, 30, original emphasis). Such omissions and under-representativeness can also be observed in the settings of Vietnam, where 'one out of two kids don't get to school' and, so, the OECD PISA is 'really just testing the elites' (OECD Interview#1, 2018). According to UIS data for the school year ending 2018, about 258 million children and youth are out of school, of which 59 million are primary school age (about 6 to 11 years old) and 62 million are lower secondary school age (about 12 to 14 years old). Indeed, these missing millions constitute a large fraction of the 'poorest of the poor' (Carr-Hill 2013) and they often reside at 'the bottom of the pyramid' (Wagner et al. 2018).

Across PISA-D countries, the proportion of the out-of-school population is also striking. Taking the 15-year-old age group as an illustration, the proportions range from 10.3% in Ecuador to 47.9% in Zambia. This means that PISA-D would fail in its aims if it did not incorporate the out-of-school youth in the assessment. Specifically, including the out-of-school youth is significant in three respects. First, if successful, it could

complement the in-school assessments and provide meaningful context for interpreting the in-school assessment results. For the OECD, by virtue of the out-of-school assessment attempted in PISA-D, 'PISA will, for the first time, be able to report on what *all* 15-year-olds in a population know and can do' (OECD 2019c, 57, original emphasis). Second, it would help to fill a data gap arising from the lack of adequate instruments to identify and reach out-of-school youth, measure the scope of and assess the reasons for exclusion, and assess the yield of the different levels of education at age 15 to inform policy and planning. Third, it would contribute to the monitoring of SDG 4, which emphasises leaving no-one behind.

Therefore, in order to address the challenge that so many 15-year-olds are out of school, PISA-D introduced the out-of-school component to incorporate this population into the assessment. The component was conducted on tablet computers and took place at the youth's home. Only five countries, except for Bhutan, Cambodia and Ecuador, took part in this activity. The main OECD report (2018c) noted that Bhutan joined PISA-D later than other countries and only conducted the cognitive test. However, in the case of Cambodia, the Ministry of Education, Youth and Sport (MoEYS) considered that the implementation of the out-of-school component would be too costly and would not have any credibility (explained further below). Considering the challenge of locating a single year age group within these countries, the out-of-school component also enlarged the target population age range from 15-year-old to 14–16-year-old. Table 6.2 reproduced below provides an overview of the resulting coverage of the out-of-school target population broken down into the six zones of exclusion.

The Educational Testing Service (ETS) was the private contractor to deliver the out-of-school component. It is the world's largest educational testing company based in the U.S. According to my interview participant, who was initially involved in this aspect of work, the ETS is 'a school testing service and has never carried out a household survey let alone one as difficult as this' (Expert Interview#1, 2018). The difficulty lies in four main areas: (i) understanding the characteristics of out-of-school 14–16-year-olds; (ii) counting and locating them; (iii) constructing a sample frame; and (iv) finding, testing and interviewing this group of population. Although all of these issues were raised in the PISA-D working paper on *Technical Strand C: Incorporating Out-of-School 15-Year-Olds in the Assessment* (published in 2015), they were not resolved, and I examine them below.

Table 6.2 Information about PISA-D out-of-school target population by zones of exclusion.

	Guatemala	Honduras	Panama	Paraguay	Senegal	Average
	%	%	%	%	%	%
Zone 1: Never enrolled	13	5	14	3	30	13
Zone 2: Primary dropouts	24	21	5	14	14	16
Zone 3: Grade 6 or below	18	17	14	20	42	22
Zone 4: Primary leavers	31	42	17	22	8	24
Zone 5: Secondary dropouts	8	11	47	29	4	20
Zone 6: Fading out	5	3	3	12	3	5
Total	100	100	100	100	100	100

Source: OECD (2020, 6)

Regarding the characteristics of out-of-school youth, PISA-D's definition of 'out-of-school' was claimed to have drawn on the work of UIS and UNICEF (2010) which defined children and youth who are excluded from basic education. Accordingly, the PISA-D out-of-school sample included youth aged 14–16 who: 'have never attended school, have attended school but dropped out of primary school, have completed primary school but did not continue to secondary school, entered secondary but dropped out, or are currently enrolled in school but are *in Grade 6 or below*' (OECD 2017b, 1, my emphasis). While PISA-D's calculation method seems to be in line with the definition given by UIS and UNICEF, the last scenario – children 'currently in school but are in Grade 6 or below' – is not justified by their approach. It even contradicts the UIS and UNICEF ideal, which focuses on the rights of children having *access* to basic education and, as such, those who are in school but below the appropriate grade for their age should not be considered as 'out-of-school'.

Nevertheless, the PISA out-of-school sample included '14–16-year-olds who are not in school and those who are in school but enrolled below the PISA target grades (i.e. Grade 6 or below)' (OECD 2017b, 1). The UIS and UNICEF approach does not address 'the problem of defining current status vis-à-vis school in a study dedicated to finding, identifying and persuading specifically 14–16-year-old out-of-school youth to take a test' (Expert Interview#1, 2018). By including the latter scenario of 14–16-year-olds enrolled in school of which the average percentage is up to 22%, ETS can enhance PISA-D's accessibility to the out-of-school target population. This is particularly true, for example, in Senegal where the group of 14–16-year-olds below 7th Grade accounts for 42% of the out-of-school target population (see Table 6.2).

Perhaps, this explains Cambodia refusal to take part in the PISA-D out-of-school component, even with 'substantial interest' in this strand of work. The country's *Capacity Building Plan* (OECD 2016a, 11) raised concerns about the suitability and utility of the PISA-D out-of-school target population definition and the implementation of this strand:

> There is substantial interest from stakeholders in the formation described by Strand C, and the proposed programme design complements the Non-Formal Education National Action Plan. However, expected difficulties and costs discouraged stakeholders from supporting the implementation of this Strand. Specifically, given the challenges with identifying a target population that was both *meaningful* for research purposes and *relevant* to national interests, the consensus within the Ministry of Education, Youth and Sports is that the implementation of Strand C in Cambodia would be too expensive and not have *sufficient utility*. (my emphasis)

The expected difficulty indicated in the *Capacity Building Plan* illustrates the unique challenges that ETS faced in counting, locating and accessing the out-of-school 14–16-year-olds in each participating country. As my interview participant explained,

> In most countries, a household survey would involve visiting about four times as many households as the number of 14–16-year-olds required, and, if an area approach is adopted, it would rather like looking for a needle in a haystack. Moreover, in terms of the particular concern – out-of-school 14–16-olds' – these estimates of the accessibility of the target population are optimistic by an average factor of three, because, of course, they include 14–16-year-olds who are in school. (Expert Interview #1, 2018)

Indeed, during the field trial, it was noted that, in Honduras and Paraguay, 'out-of-school youth were twice as hard to identify, approach and interview than they were in other countries' (OECD 2019c, 63).

Due to the rarity of the out-of-school target population (except for those enrolled below PISA target grades), to complete a large sample size seemed unrealistic. While choosing fewer sampling points and fewer respondents may be plausible, it would make the sample less reliable as a basis for generating national estimates. The issue was not resolved in the PISA-D sampling frame, which relied mostly on a probability sample. That means, 'PISA-D randomly selects households and trained interviewees go to each household to apply a screener questionnaire to confirm if there are any eligible youth' (OECD 2017b, 2). This procedure was further complemented with non-probability sample with referrals from schools or interviewed youth, or household members in households with no eligible youth. The non-probability sample techniques were used for 'purposive selection of low-cost samples' (OECD 2019c, 59). In these ways, important minorities and marginalised groups were very likely under-sampled.

> Determining the weights to put on such a combination of selected areas in terms of their relative likelihood of including out-of-school 14–16-year-olds in order to produce final estimates would be complicated guesswork, not science. (Expert Interview #1, 2018)

Therefore, although the OECD (2017b, 2) suggested that the sampling of out-of-school youth would serve to 'yield a large enough sample' to test the 'validity' of the cognitive test and contextual questionnaires and to link to PISA and PISA-D school-based assessments, this can be hardly achieved. To establish the linkage, a more complex chained two-stage approach would have to be applied: first, to link the country-specific Strand C instruments to the school-based ones and, second, to link the national country-specific instruments to the main PISA. For this to happen, it has to assume that there are

sufficient common items in each linkage pair, which would not always be possible. Moreover,

> PISA frameworks for the topics covered in the contextual questionnaire are intrinsically built around the concept of learning in school but there is no reason to conceptualise the out-of-school 15-year-olds in terms of school-based concepts. (Expert Interview #1, 2018)

These issues were implied in the PISA-D *Project Completion Report* on the outputs from the field trial (FT) of the out-of-school component,

> the sample sizes across all participating countries were smaller than planned but were adequate for finalising for the Main Survey [MS] instruments and design. … However, a lack of sufficient data at the *item level* across countries prevented the full set of FT analyses to be completed as planned. The analyses intended for the FT were postponed to the MS phase and contractors noted there was *uncertainty* about the *appropriateness* of the items for the Strand C population that would not be fully understood until after the MS analyses are completed. (2019c, 61, my emphasis)

Essentially, I am suggesting that the out-of-school component was unable to access a sufficient number of out-of-school youth. The OECD and ETS resolved this problem by redefining such youth to include those who are in school and have repeated a grade, as they were more readily accessible. The definition employed has no similarity to that used by other agencies or studies which ensures that the data cannot be compared to previous studies.

Nevertheless, the OECD concluded that the field trial was successful and that the outputs of the field trial 'are a valuable contribution towards the design of the Main Survey, and for accuracy of the data collection and submission practices in the out-of-school component' (ibid.). At the PISA-D International Seminar (2019d, 5), however, Ward stated that 'this is effectively guesswork and is carried out under assumptions that are not underpinned by real evidence', but since 'in the context of SDG 4, there is no substitute for assessing the skills of the whole population […] this is the challenge that the out-of-school component of PISA-D sought to overcome'. As I have shown above, PISA-D has not been able to resolve this bottleneck, but for Ward, it has succeeded not because of its technical validity and reliability but because there is no 'substitute' – it is 'in demand'. The OECD's objective was to mainstream 'the out-of-school assessment in future cycles of PISA' (ibid.).

Therefore, the experimental success of PISA-D was preordained. The status of PISA-D as a pilot project, rather than to generate real evidence of effectiveness, has primarily served to get PISA-D off the ground (anticipating scholarly critiques and scepticism from countries and other agencies). Piloting PISA-D also helps to demonstrate the OECD

commitment and responsiveness to the unique challenge that the international community faced to assess 'the skills of the whole population'. From the perspective of organisational legitimacy, displaying such responsiveness is often more pragmatic (and easier) than producing immediate results (Suchman 1995). Given that the success of PISA-D is largely socially constructed, the process relies heavily on strategic communication and environment manipulation (see also Chapters 5 and 7).

In this section, through carefully examining the technical procedures involved to achieve the stated objective that PISA-D will be 'contextualised' for low- and middle-income countries while maintaining comparability with the main PISA, I have revealed and questioned the outputs achieved and the assumptions implicated in the instruments. My interrogation suggests that the PISA-D exercise subscribes to the 'Anglo-Saxon testology', which rushes to categorise and is 'nothing but a ridiculous simplification of knowledge and a robbing of meaning from individual histories' (Malaguzzi [1990], cited in Moss 2016, 173). In what follows, I analyse the findings generated from the PISA-D school-based assessments[10] and the OECD's application of the EPA to make sense of these findings.

Findings from PISA-D school-based assessments

The findings of PISA-D are presented in both the main OECD report and individual country reports. The individual country reports cover fine-grained analyses, whereas the main report – *PISA-D Results in Focus* – provides a comparative overview of the results across PISA-D countries. The main report also compares students' performance, where possible, to the OECD average for the PISA 2015 assessment cycle in order to establish linkage to the main PISA. As noted, the educational prosperity framework was applied to help frame and interpret the findings presented in the reports. Thus, in the main report, the findings sequentially cover educational attainment at age 15, students' performance in reading, mathematics and science, variation in student performance, and health, well-being and attitudes towards school and learning. These are structured in accordance with

---

[10] The OECD released the final report on the results of the out-of-school component in December 2020. Since I had completed the thesis by that time, my analysis has not been able to incorporate the findings from this report, apart from the table which covers the percentage of the target population of 14–16-year-olds in each country.

the 'prosperity outcomes' identified within the EPA. In this way, the OECD can interpret these 'outcomes' through connecting to the contextual factors that are referred to as the 'foundations for success' (see Figure 6.2 for 'the Prosperity Tree').

In summarising educational attainment at age 15, the main report contrasts the average percentage of all 15-year-olds enrolled in Grade 7 or above in PISA-D countries with the OECD average (i.e. 43% vs. 89%), before breaking this figure down with regard to the proportion of boys and girls who fulfilled the requirement. It then notes that only 62% of students who sat the PISA-D test were enrolled in the expected grade or in a higher grade given their age, and that many of the students who had fallen behind indicated that they had repeated a grade in primary or lower secondary school. Whereas the OECD average for students reporting having repeated a grade is 12%, grade repetition in PISA-D countries ranges from 18% (in Ecuador) to 50% (in Senegal). Obviously, such differences would require different approaches to address countries' needs, rather than standard procedures. The differences also suggest that it is not meaningful to compare the results with the OECD average and even across PISA-D countries. In this regard, PISA-D replicates, rather than addresses, the more fundamental issues concerning schooling in low- and middle-income countries.

The results of PISA-D school-based assessments in reading, mathematics and science are reported through three major modes of comparison. The first involves an extended commentary on how the seven PISA-D countries compared with the OECD average in PISA 2015. The standard in this mode of comparison is the minimum level of proficiency envisaged in SDG 4 (benchmarked as PISA level 2). For example, the report (2018c, 6–7) stated that 'about 23% of students across PISA-D countries achieve the minimum proficiency level in reading, compared to the OECD average of 80%', 'about 12% of students across PISA-D countries achieve the minimum level of proficiency in mathematics, compared with the OECD average of 77%', and 'about 18% of students across PISA-D countries achieve the minimum level of proficiency (Level 2) in science, compared with the OECD average of 79%'. Table 6.3, which is reproduced below, provides a 'snapshot of performance' across reading, mathematics and science, with the percentage of students achieving the minimum proficiency level in each discipline and these are compared to the OECD average as well as the lower-middle income average.

Table 6.3 Snapshot of performance in reading, mathematics and science.

	Mean reading score	Mean mathematics score	Mean science score	Coverage of the national 15-year-old population (PISA Coverage index 3)	Education Sustainable Development Goal (SDG 4) indicator	
					Students achieving minimum level of proficiency (Level 2) in reading	Students achieving minimum level of proficiency (Level 2) in mathematics
	Mean	Mean	Mean	%	%	%
Cambodia	321	325	330	28.1	7.5	9.9
Ecuador	409	377	399	60.6	49.4	29.1
Guatemala	369	334	365	47.5	29.9	10.6
Honduras	371	343	370	41.4	29.7	15.4
Paraguay[1]	370	326	358	m	32.2	8.3
Senegal	306	304	309	29.0	8.7	7.7
Zambia	275	258	309	36.0	5.0	2.3
OECD average	493	490	493	89.0	79.9	76.6
Lower-middle income average	378	368	392	60.2	37.7	28.7

Source: OECD (2018c, 7)

The second mode involves comparing some of the PISA-D countries with countries that have already participated in PISA. The intention here is to relate PISA-D countries to the existing members of the global PISA community, such as Brazil. An example of this can be seen from the extract: 'In Ecuador, 49% of students reached this level [Level 2] – the highest percentage among PISA-D countries, and a similar percentage to that in Brazil in PISA 2015' (ibid., 6). Given that the overall low levels of performance across PISA-D countries (even lower than the lower-middle income average), this mode of comparison appears difficult. However, as the analyses of the country reports will show, this type of comparison serves as a prominent approach to the promotion of the transfer of 'effective' policies to PISA-D countries.

Third, in addition to commenting on the performance of PISA-D countries against the minimum level of proficiency (i.e. Level 2), the report further breaks down each individual country's performance in accordance with the lower proficiency levels newly added to the PISA development scale (such as Level 1a, 1b and 1c). For instance, 'in Ecuador, Guatemala, Honduras and Paraguay, Level 1a was the median reading proficiency level among students. This means that more than half of the students could, at best, retrieve pieces of information that are explicitly stated'. Also, 'in Cambodia, Senegal and Zambia, the median reading proficiency among students was Level 1b, meaning that more than half of the students could, as best, solve the easiest text-comprehension tasks included in the PISA-D assessment'. Finally, 'in Zambia, Level 1c was the median mathematics proficiency level. Most students in this country could only understand mathematics questions involving simple, everyday contexts where all relevant information is clearly given and defined in a very short simple text' (ibid., 6–7).

With extensive coverage on how poorly the majority of students in PISA-D countries performed, the report also noted that 7% of students, on average, 'demonstrated high levels of knowledge and skills in reading, meaning that they scored at or above Level 3 – the *typical* level of proficiency among 15-year-old students in OECD countries' (ibid., my emphasis). It then suggested that PISA-D countries should aim for increasing 'the share of their students who perform at these higher levels' (ibid.). The explicit message is, thus, that PISA-D countries should follow and catch up with the OECD countries by improving their PISA performance. Under this logic, levels of development are seen only through the prism of PISA, and any improvement will be demonstrated through a positive change in their PISA scores. By doing so, the OECD successfully brought all countries

into one common scale in line with PISA, within which PISA-D countries – located remote in space and time – are encouraged to begin their ascent. It is noticeable here that the 'other' is elaborated through degrees of deviation in relation to the OECD average or to the performance of some exemplars (such as Brazil), but underlining these are the standards and path pre-defined by the OECD (e.g. Level 1a, 1b and 1c).

Having classified the varying performance levels, the report turned to address the issue of equity by exploring variation in student performance. Equity in this report refers to that 'all children have access to education opportunities that lead to quality learning outcomes, irrespective of their gender, their ethnicity, or their parents' wealth, education or occupation' (ibid., 8). Accordingly, the report drew on information about students' backgrounds, while acknowledging that this only offers a partial description of inequalities and unfair opportunities, given that the current report does not cover the findings from the out-of-school component. It, nonetheless, provides a series of shallow observations. For example, girls tended to outperform boys in reading, but the gap was less pronounced than in OECD countries; in mathematics and science, gender gaps favoured boys and girls differently across countries and with varying levels of significance; students in urban settings outperformed rural counterparts; the mean performance of students at different levels of socio-economic status showed that students in PISA-D countries tended to do worse than students with similar socio-economic resources in OECD countries. These are generic statements, which appear to confirm expected performance levels relative to the OECD countries. The lack of information about those (i.e. 57%) who are not included in the school-based samples suggests not only the superficiality and simplification, but, also, undermines the value of PISA-D.

There is a further problem that effects the findings from the school-based assessments: the proportion of students who do not speak the language of instruction at home. In the report, it was noted that 'a significant minority of students' in Guatemala (9%) and Paraguay (41%) do not speak the language of instruction (Spanish) at home and that 'the vast majority of students' in Senegal (94%) and Zambia (83%) do not speak the language of instruction (French and English respectively) at home. While the issue was recognised in the PISA-D documents, the focus was primarily on the collection of information on this, exemplified by the contextual questionnaires collecting 'more detailed data on students' language of instruction at school, language at home and their socio-economic status, as measured by home possessions and parents' education, literacy skills and

participation in the labor force' (OECD 2018c, 19). The issue, however, was not addressed in terms of how it would impact the validity of the data generated from the cognitive test on students' performance. The main report only noted that 'students who speak the language of instruction at home scored significantly higher in reading than students who speak a different language at home' (ibid., 9). The issue of language raises fundamental questions as to the reliability of using data to compare education outcomes and identify policy lessons across the seven countries, or with reference to PISA more generally.

When trying to link the performance and findings to contextual factors, the report first rehearsed the explanatory logic of the EPA: that students' performance on the PISA-D tests at age 15 is 'the result of an accumulation of various factors that affect children's development, beginning at conception and continuing through to the time of assessment' (ibid., 11). The report thus issued a caution against establishing a strong relationship between student performance and school or classroom practices. Despite this, it suggested that it is still possible to identify a range of factors that influence students' performance and the related outcomes. These factors were referred to as the 'foundations for success', which include resources, inclusive environments, learning time, quality instruction, and family and community support. Although it is wholly unclear where such evidence was, the report claimed that 'PISA-D provides evidence of how these factors are related to 15-year-old students' performance' (ibid.). In the following section of the report, the findings were, thus, elaborated through relating to each of these 'foundations'.

Specifically, most of the space was devoted to resources invested in education, with the finding that high spending is associated with higher student scores in the cognitive domains tested. It drew reflections across a range of factors, such as student-teacher ratio, teacher training and professionalism, physical condition of school infrastructure and the availability of textbooks. In particular, the student-teacher ratio across PISA-D countries was contrasted to the OECD average (i.e. 13 students per teacher) and to that in the existing members of the global PISA community, such as, from 'almost 30 students per teacher in Brazil, Colombia, the Dominican Republic and Mexico, to fewer than 10 students per teacher in Albania, Belgium, Greece, Hungary, Luxemburg, Malta and Poland' (ibid., 12). Regarding inclusive environments, the focus was on students' sense of belonging to school, along with their safety conditions both at school and on their way to or from school. Surprisingly, the report noted that 88% of students across PISA-D

countries felt that they belonged at school, which out-performed the OECD average of 73%.

In terms of learning time, students across PISA-D countries reported that they had skipped a day of school recently (33%), that they had missed school for more than three months in a row (15%–24%) and that teachers were frequently absent; these were reported at higher rates than the OECD countries. As for the quality of instruction, the findings were mainly concerned with student-teacher relations, teacher effectiveness and class environment. For example, most students reported positive views about their teachers, but there was noise and disorder in most lessons, which was suggested as an indication of teachers' effectiveness. Finally, to refer to the wider learning environment of families and communities, the questionnaires asked how often parents or other family members engaged in exchanges and activities with students. This was then presumed to serve as an indication of the family support for student's engagement at school and with learning. Although teachers reported that only 28% of students in PISA-D countries were in schools where parents often or always attend parent-teacher meetings, most students reported that their parents encouraged them to get good grades and emphasised the importance of completing secondary school.

By and large, the findings examined above illustrate how the report tries to link students' performance to the contextual factors by applying the five 'foundations for success' prescribed by the EPA. However, given that the EPA emphasises the cumulative and interrelated effects of these 'foundational factors', it is unclear how the relationships can be established. Therefore, the grouping of findings in relation to these 'foundations' are rather arbitrary and offer little policy guidance. Moreover, it is not meaningful to compare the 'foundations' within PISA-D countries (already recognised as low-resourced settings) to those established within countries that are more affluent (such as Albania, Belgium, Greece, Hungary, Luxemburg, Malta and Poland).

Below, I examine the 'effective' policy interventions generated from these findings that are presented in both the main OECD report and individual country reports.

Policy recommendations

Effective interventions in the main report

> PISA-D assessment results provide countries with a solid database that can help them refine policy priorities and set new goals and targets to improve the foundations for success at all levels of their education systems. The data collected have a lot to say about the allocation of resources and its implications for equity. With reliable data on differences in outcomes and gaps in access to the foundations for success between groups of children and young people, countries can determine whether poor and marginalised populations are given equal opportunities to succeed at school and beyond. The challenge for countries overtime is to maintain a focus on these goals and targets, and to track progress towards them by participating in future cycles of PISA. (OECD 2018c, 16)

As these statements show, the OECD claimed that the findings from school-based assessments provide countries with the data and evidence to help refine policy priorities and to improve their education systems. In particular, they help to inform 'the allocation of resources' and 'its implications for equity' within countries (ibid.). Thus, one policy lesson presented in the main report is that the findings tend to confirm the success of PISA-D countries' efforts to implement gender-equality policies over the past decade. For example, 'gender differences in attainment and proficiency are not large and are sometimes in favour of girls across PISA-D countries' (ibid.).

This policy lesson, however, contradicts the challenges identified in the PISA-D background working paper on Technical Strand C, which suggested that 'gender', along with 'ethno/religious group' and 'urban/rural location', are the main structural features affecting out-of-school youth within PISA-D countries. More specifically, it identified that, in all of the countries, 'in general, girls are more likely than boys to be out-of-school' (OECD 2015, 18). The reasons for these include, for example, schools are not available in the local area, or parents believe that access to the schools that are available is insecure for girls, or that the community may not support girls' education. However, these issues were not reflected in the findings, simply because the school-based assessments did not cover the most disadvantaged and poorest rural girls, as my interview participant pointed out.

The other policy actions recommended in the main report are similarly controversial. For instance, it identified policies that had a negative effect on student performance and, therefore, suggested that they need to be revised. One of these is the low standards for hiring teachers. The report suggested that by raising standards, countries can 'set a course towards high-quality instruction and classroom management, the effects of which will gradually contribute to improved learning outcomes' (ibid., 16). As the findings on the quality of instruction above have shown, this claim was mainly derived from observations

that, on average, about 29% of students reported noise and disorder in most lessons and 25% reported that they do not start working for a long time after the lesson begins. These answers were perceived as the indicators of teacher effectiveness.

However, it is unclear how these findings led to the conclusion that PISA-D countries should prioritise the recruitment of teachers. Arguably, these could be more closely related to large class sizes: in Cambodia and Senegal, the teacher-student ratio is 1:30, whereas in Zambia, the ratio is 1:43. This policy lesson of raising standards is an apparent anomaly in the low-resourced settings where teachers are in short supply and where salaries are low and teachers work in additional jobs. For example, in Cambodia and Senegal, 'more than half of all teachers who responded to the questionnaire, particularly those in urban regions and those who work in private schools, work as private tutors in addition to teaching at school'; in Cambodia and Paraguay, 'more than one in three teachers also work in another job that is not related to teaching' (ibid., 12).

Despite the problem identified, the main report suggested that countries should aim for establishing in their education systems the 'foundations for success' and should track progress towards them by 'participating in future cycles of PISA' (ibid., 16). The specific policy interventions proposed as 'effective' for achieving this goal is summarised in Table 6.4. As the table shows, the policy interventions promoted in the main report are, rather, generic and decontextualised. They do not offer a critical and deeper reflection of the realities of PISA-D countries. And many of the interventions – framed by the EPA – are conceptualised in terms of school-based concepts (e.g. quality, efficiency, competency, diversity, evaluation, and teacher professionalism).

Among these 'effective interventions', the most repeated and detailed one is *grade repetition*. The report suggested that grade repetition is a costly policy, 'as it requires greater expenditure on education and can delay students' entry into the labour market'. While it also noted that 'research has found mainly negative effects of grade repetition on academic achievement and attainment' (ibid., 16), there was no evidence showing how the PISA-D results serve to demonstrate that improved outcomes are driven by the elimination of grade repetition rather than by many other foundational factors that the EPA identifies.

Moreover, PISA-D data reported on levels of student performance and the proportion of students having repeated a grade provide very limited support for that claim. For example, grade repetition in Senegal (55.5%) is higher than that in Zambia (43.3%) but

student performance in reading and mathematics in Senegal (306 and 304, respectively) are higher than performance in Zambia (275 and 258, respectively). Similarly, grade repetition in Guatemala (36%) is higher than grade repetition in Cambodia (29.1%), but student performance in Guatemala in all subjects (369, 334, 365) is higher than that in Cambodia (321; 325; 330). Equally, grade repetition across PISA-D countries, on average (32.1%), is higher than that in Cambodia but the PISA-D averages in reading and science (346 and 349 respectively) are consistently higher than those in Cambodia (321 and 330 respectively). Table 6.4 provides an overview of grade repetition rates across PISA-D countries and others.

Figure 6.4 Grade repetition rates across PISA-D countries and others.

Percentage of students who had repeated a grade in primary, lower secondary or upper secondary school
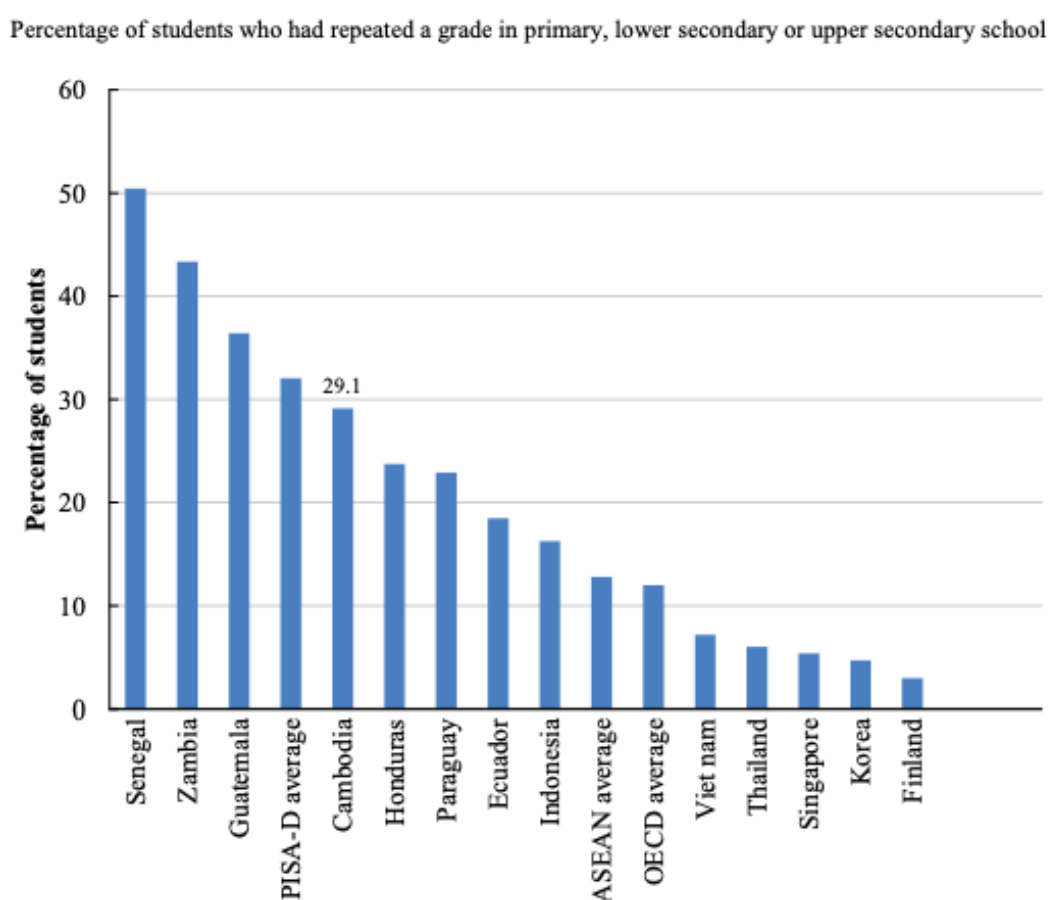
On 29 September 2020, the OECD launched the PISA 2018 (Volume V) report, entitled *Effective Policies, Successful Schools*. Schleicher presented the latest findings, focusing on issues relating to school organisation and the policies and practices that define how education systems work and change over time. When he was asked whether there

are policies of particular countries that he thinks affect grade repetition one way or another, he gave the answer that

> grade repetition is a very *convenient* and *easy to deploy* mechanism to deal with underperformance. The other ones, you know, providing additional support, providing extra lessons, are much more *complex to manage* for schools … In some countries […], schools are only serious if there are consequences for underperforming students, however, what we can show, I think, quite clearly that none of those policies [sic]. There is no country where grade repetition really is of use to better learning outcomes or more equitable distribution of outcomes. (Schleicher 2020, my emphasis)

Therefore, beneath the EPA's humanitarian framing, we can still observe the economics of education approach underpinning PISA, which are reflected in the concern regarding students' timely entry into the labour market and the efficiency of public spending.

Table 6.4 Effective interventions based on findings from PISA-D school-based assessments.

Establish the foundations	Effective interventions

Better allocation of resources	**Eliminate policies that are costly and have no positive impact.**
	  Costly policy: grade repetition *Alternative*: offer support to struggling students
	Interventions:
	  a. remedial classes during regular school days & in-term breaks.
	  b. providing additional learning time for differentiated instruction.
	
	Allocate resources more equitably across schools.
	  Interventions:
	  a. give more resources to schools in rural, disadvantaged and/or public schools.
	  b. eliminate practices that undermine the equity of education systems.
	    e.g. improve the quality, not the quantity, of teachers in disadvantaged schools.

Adequate learning time	**Reduce student truancy and teacher absenteeism.**
	  Interventions:
	  a. consider the reasons for student truancy and teacher absenteeism.
	  b. make better use of the allocated time and resources for learning.

Improve school environment	**Provide an environment that is conducive to learning.**
	  Students: regular attendance; disciplined and respectful.
	  Teachers: co-operate and exchange ideas or material; supportive.
	  Principals: provide learning opportunities to students with special needs.
	      reactive to behavioural & academic problems.
	      offer extracurricular activities at school.
	  Parents: participate in school activities; interact with other parents.
	  Governments: use assessments, information systems & informal mechanisms.
	      identify struggling schools.

	Ensure that education set the tone for a caring and inclusive school community.
	 Principals and teachers:
	    a. address the diversity of learners.
	    b. respond to special needs of students.
	    c. responsible for educating all children.

Improve the quality of instruction	**Make teaching more effective.**
	 Ensure competent people want to work as teachers.
	 Ensure high-quality teaching.
	 Interventions:
	    a. devise a system of motivating, attracting, developing and retaining high-quality teachers & school leaders.
	    b. strengthen pre-service training requirements & improve teacher preparation.
	    c. evaluate teaching and provide feedback for improvement.
	    d. align professional development with teaching goals, and instructional resources with a competency-based curriculum.

Strengthen family & community support	**Schools help parents overcome barriers to engage in child's school life & learning activities.** **Intervention:** a. School leaders ensure that parents recognise the importance of school and learning.

Policy options across country reports

PISA-D countries did not produce the country reports on their own. The OECD and its private contractors have provided training and assistance in data analysis, interpretation, report writing and communication. More specifically, each country formed an analysis task force and nominated a lead analyst on data analysis and reporting. The OECD worked directly with the analysis task force, especially the lead analyst from each country. Indeed, the OECD carried out a lead analyst programme from September 2017 through December 2018. The aim of this programme was to help lead analysts to complete their analysis of PISA-D data and prepare a national report. The details of the programme can be viewed in Table 6.5.

Table 6.5 PISA-D lead analyst programme.

Timeframe	Learning Targets
Sep-Dec 2017	Introduction to PISA data and its uses for policy makers, educators, researchers; guided reading
Jan-May 2018	Preparing for analysis of main study data: understanding the PISA design; limitations of PISA samples; understanding and interpreting PISA coverage indices; the analytical framework
	Drafting of introductory Chapter 1 of national report
May-Jul 2018	Analysis of data and drafting of Chapters 2–6 of national report: training in data-analysis techniques and methodologies, collaborative drafting with OECD (residence at OECD, Paris)
Jul-Nov 2018	Finalising the national reports and preparing for launches, including producing country-specific dissemination materials (PPTs, country notes, Tweets, videos, webpages, etc.)
Dec 2018	Publication and dissemination of national reports with final results from the school-based assessment

Source: PISA-D Brief – 25 (OECD 2018e, 2)

Understanding the OECD's role in the process of national data analysis and reporting provides an important context for analysing the structure and contents of these national reports. According to the OECD (2019c), there had been two phases scheduled for the lead analysts when they resided at the OECD Headquarters in Paris from May to July 2018. The first phase (May) focused on cultivating statistical skills, including understanding 'the characteristics of international assessments and examples of research conducted using their initial PISA-D data files', 'how to handle international studies' databases and technical documents', 'how to replicate PISA results and answer research questions using relevant software (such as Mplus, R, STATA, SPSS or SAS)', and

'practical experience on how to formulate, analyse, interpret and present results relevant to educational research using PISA-D data' (ibid., 72). In the second phase (June–July), the lead analysts 'applied the skills acquired through the training to analysis of their own PISA-D data and, with the mentoring and support of the OECD's PISA-D manager, PISA Analysts and Experts, interpreted their results and prepared a report on the basis of the OECD's national report template' (ibid.).

Therefore, each report, containing six chapters, is based on a template prepared by the OECD, which includes analyses of the four 'prosperity outcomes' and the 'foundations for success' that underpin them. Although the template was not provided on the OECD website, based on the table of contents and chapter outlines given in each country report, I extracted and summarised the template in Table 6.6.

As the template shows, the structure and content of the national reports have been strongly influenced by the EPA, so much so that the identification of contextual factors that can be related to student performance is not derived from the actual data but based on the logic already established within the framework. In other words, the interpretation of results serves to stabilise the EPA, especially the assumptions about the relationship between 'prosperity outcomes' and 'foundations for success'. From this, the template not only exemplifies the singularity of perspective, but also contributes to legitimating the authority of the OECD and its private contractors.

Furthermore, based on the EPA's life-course approach, students' performance on PISA-D at age 15 is the result of an accumulation of those interrelated foundational factors from the conception stage. In this regard, it is unfeasible to extrapolate 'school effects' from the results and to suggest specific policy interventions. Nevertheless, as can be viewed from the template, the reports devoted Chapter 6 to this, highlighting policy options that countries could adopt in order to move towards establishing the foundations and thereby improving their educational outcomes. Table 6.7 provides a summary of these policies, outlining the main problems identified, the key policy actions recommended, and the sources of evidence cited to support those policies.

Table 6.6 Summary of the OECD's national report template.

Chapter 1. Country (X) in PISA-D

The chapter firstly introduces PISA and PISA-D and why the country is participating. It then explains how the information collected in the assessment can be used to compare the education systems in country (X) to other countries. The analytical framework draws on the EPA, in particular, the 'prosperity outcomes' and 'foundations for success' which correspond to the 5th stage of child development (i.e. late primary and lower secondary; ages 10–15).

Chapter 2. Achievement and attainment outcomes at age 15 in country (X)

The chapter discusses the PISA-D results and what they reveal about the achievement and attainment outcomes in the country. It looks at the enrolment of 15-year-olds and their attainment, paying particular attention to the role of grade repetition. This information is considered as providing important background for understanding student achievement in the key subjects, and for comparing the country's performance with other countries.

Chapter 3. Health, well-being and attitudes toward school and learning at age 15 in country (X)

The chapter describes the health, well-being and attitudes toward school and learning of 15-year-olds in the country. It identifies for each outcome a vulnerable population with poor health and well-being or negative attitudes towards school and learning, and summarises the relationship with the achievement outcomes that are discussed in Chapter 2.

Chapter 4. Foundations for success in country (X): Resources invested in education

The chapter examines the resources invested in education in the country and makes comparisons with other PISA and PISA-D participating countries and economies. It then explores how these resources are allocated across schools, and the relationship between educational resources (financial, material and human resources) and student performance.

Chapter 5. Foundations for success in country (X): The school and community environment

The chapter analyses the four aspects of the foundational factors, including inclusive environment, learning time, quality instruction, and family and local community support. The chapter describes, in particular, how the presence of these foundations varies between students and schools in the country.

Chapter 6. Looking forward: Policy Options for country (X)

The chapter first summarises key findings of PISA-D, emphasising both the 'prosperity outcomes' and 'foundation for success' of education in the country. The chapter then present policy options that can have both short-, medium- and long-term impacts on education development in the country.

Table 6.7 PISA-D national reports: summary of problems, recommendations and evidence.

Country	Problems (identified/prioritised)	Key policy recommendations (Short- and medium-term improvements)	Sources of reference/evidence (countries and research/reports cited to support the policies advocated)
Cambodia	Low investment in education	*Low hanging fruits or quick wins:* *(policies that do not cost much but still can have impact)*	**PISA 2015** suggests that the right education likely to compensate for the low investment in education. e.g. **Vietnam** has lower spending per student but outperformed many OECD countries in science; **South Korea** spends well below the OECD average per student but becomes the highest-performing OECD country in mathematics; **Shanghai**'s 10% most disadvantaged 15-year-olds perform better in math than the 10% most privileged students in the **US** and several **European** countries.
	Grade repetition	*to reduce grade repetition* *(1) provide remedial classes or tutoring during or before semester breaks.* *(2) provide extra teaching time or practices.* *(3) ensure students start school on time or are on track.*	**OECD (2013; 2016b)** shows grade repetition is not helpful but even harmful to student learning (esp. **French** experience in PISA). **OECD (2013):** Grade prepetition can be a costly policy, as it requires greater expenditure on education and delays students' entry into the labour market. **PISA 2015** results show high-performing nations (e.g. **Japan, Chinese Taipei, Vietnam, Finland, Estonia, Singapore & Canada**) have minimal grade repetition rate. **UNESCO (2010):** If resources spent on repeating a grade were spent on enrolling new students into school without reducing the quality of education, annual GDP globally would increase by 0.37%; the growth rate was in particular larger in low-income countries (UNESCO, 2012). **PISA-D results** show grade repetition is associated with lower levels of academic performances. **UNESCO (2012):** Creating extra classes, peer learning activities or study clubs is more cost-effective than giving them a second chance to repeat a grade. **France** experience in PISA shows grade repetition does no justice to students. **OECD (2016b):** Reducing grade repetition with support from school improves students' performance in PISA accordingly.
	Far from reaching the international bar	*to ensure quality learning time* *(1) strengthen school discipline.* *(2) build a strong school management system.* *(3) make learning assessment serious.* *(4) supplement students with extra competency-based practices e.g. homework or routine tasks in the classroom.*	**OECD (2013; 2016a):** The positive effect of school discipline on student performance is clearly evident in many PISA countries e.g. **Vietnam, Korea, Japan and B-S-J-G (China)**. **Previous PISA results** consistently show when learning becomes a serious goal in schooling, it can compensate for passive teaching or even social disadvantage, as evident in **East Asian countries** such as **B-S-J-G-Macao-HK (China), Chinese Taipei, Korea** and **Japan** who are among **top-performers in science in PISA 2015**. **OECD (2013; 2016a; 2016b):** It is not about how much resources are spent but about how those resources are used, e.g. **U.S., UK, Australia & other OECD countries** spend significantly more but lag behind **Korea, Finland** and even **New Zealand** in reading.
	Absenteeism & tardiness	*to make learning visible in the classroom* *(1) abolish cheating in testing or assessment.* *(2) build a strong learning culture.*	**PISA-D results** about students' higher sense of belonging and positive values towards school and learning indicate that this is the dividend to ensure that schooling is not only a luxury but also can translate into learning. **Previous PISA results** show making learning a serious goal can compensate for passive teaching or even social disadvantage. **PISA 2015** show: in East Asian countries e.g. **Beijing-Shanghai-Jiangsu-Guangdong-Macao-Hong Kong (China), Chinese Taipei, South Korea and Japan**, the deep-rooted learning culture play a key role in their success. **OECD (2013; 2016a; 2016b)** show that in these East Asian countries, it is not about how much resources are spent but about how resources are used. **U.S., UK, Australia** and the **other OECD countries**, spend significantly more but lag behind South Korea and even New Zealand in reading performance.
	Low proficiency level	*to emphasise classroom assessment* *(1) monitor and shape what and how students learn.* *(2) promote performance & competency-based classroom assessments.*	**PISA-D results** show that about 90% of 15-year-old students in Cambodia have a low proficiency level, suggesting classroom learning needs reinforcements.
	Lowest spending per student & instructional time	*to align learning opportunities with PISA test items* *(1) increase learning tasks per lesson/chapter.* *(2) align them with competency-based curriculum or PISA test items.*	**OECD (2013; 2014):** The successful experience of **Asian countries** e.g. **Korea, HK, China Macao, Vietnam,** and **Taiwan** through complementary education indicates that this can be done.

Cambodia	**Schools are poorly resourced**	***Long-terms must haves:*** *(policies and priority actions have impact in the long run)* to ***strengthen "school standards" and use this to benchmark allocation of budgets & human resources*** *(1) create an average school standard for rural schools.* *(2) network rural schools with district- or province-based schools.*	**Shaefffer and Heng (2016):** schools can learn from schools. **UNESCO's Global Monitoring Report 2018** stresses that monitoring is the key to school development and education quality improvement (UNESCO 2017).
		to ***improve universal basic skills*** *(1) invest more in basic education.* *(2) dedicate to literacy & numeracy skills.* *(3) increase instructional hours or expand school days to promote practices of PISA cognitive skills for low grades.* *(4) incorporate competency-based teaching & learning and align with classroom assessment;*	**PISA-D results** suggest that poor basic skills (functional literacy) in the lower grades is what truly matters to the overall performance in secondary education since PISA-D measures the cognitive skills students develop across ages. **OECD (2015) projection:** increasing average achievement in current students by 25 PISA score-points has a uniform effect on all countries' GDP by 30% over the next 80 years if there is a 100% enrollment (**Hanushek & Woessmann, 2015**). **Previous evidence** consistently shows that increasing the quality of school has a large impact on economic gains.
	Low attention to teacher professionalism	to ***improve the quality of instruction*** *(1) strengthen competency-based curriculum in teacher education.* *(2) develop incentive system.*	**PISA (2012; 2015):** better teachers can compensate for the school disadvantage. In **top-performing PISA countries**, investing in the quality of teachers makes a different impact on student performance. **Schleicher (2018): successful countries** experience tradeoff between having more teachers as a response to reducing class size and good teachers by investing in competitive salaries, ongoing professional development and a balance in working time show that investment in the latter is what matters the most e.g. **Korea, Finland & Vietnam**.
Zambia	**Existing policies/programmes not sufficient to establish fully the foundations for success**	***Quick wins*** *(policies do not cost much but can have a positive impact)*	
	Low parental involvement	to ***strengthen family and community support*** *(1) implement the homework policy.*	
		Long term policies *that can have impact in the long run*	
	Low budgetary allocation	to i***mprove resource allocation*** *(particularly for rural and disadvantaged schools)* *(1) use the existing minimum school standards.* *(2) lobby for increased budgetary allocation;*	**Senegal** has allocated about 23% of total public expenditure towards education.
	Disadvantaged & rural schools: - larger teacher-pupil ratio - large number of novice teachers	to ***strengthen teacher recruitment policy*** *(1) make working in rural areas attractive.* *(2) strengthen teacher transfer policy.*	

Senegal	**Low attainment**	*to **decentralize the textbook procurement** to districts & schools* *to **reduce the school entry age to 6 years***	**UNICEF (2013)**: some students, esp. in rural areas, enter schooling when they are older than 7 years. **The National Assessment findings and examinations results** have consistently shown a relationship between age and performance with younger candidates performing better than the elderly ones. All **PISA-D participating countries** and the **OECD countries** have primary school entry age of below 7 years **(OECD, 2015).** Grade prepetition can be a costly policy, as it requires greater expenditure on education and delays students' entry into the labour market **(OECD, 2013)**.
	Grade repetition	*to **enhance the monitoring and evaluation of formative assessments*** *(1) conduct regular formative assessments to enhance remediation and reduce grade repetition;*	
	Low achievement	*to **investment more in early childhood education*** *(1) invest in teacher quality, school materials & instructional resources;* *to **strengthen teacher training*** *(1) enhance monitoring mechanism.* *(2) standardise curriculum with similar content coverage.* *(3) standardise teacher training processes.*	
	Learning is far below expectation	*to **strengthen competency-based curriculum*** *(1) revisit the orientation of school programmes & the aims of educational actions.*	The management of learning time is crucial for countries that have performed well.
		*to **improve pedagogical approaches*** *(1) make teaching engaging.*	
		*to **adapt teacher training modules*** *(1) place teachers as one of the contextual priorities.* *(2) ensure appropriate mechanisms to maintain teachers' enthusiasm & excellence.* *(3) protect and provide moral support for students.*	
	Poor teaching environment	*to **ensure teaching environment*** *(1) restore discipline and review regulations of schools.* *(2) secure school space and protect students from attacks.* *(3) provide staff with modern work tools and equipment.* *(4) encourage local community support.* *(5) support school projects developing complementary activities.*	**PISA-D results** show 31.2% of students reported noise or disorder in most lessons and 27.4% reported students do not listen; 12.6% reported they do not feel safe in school and 11% reported they engaged in physical confrontations during the last two months in school. **PISA-D results** reveal many limitations in state interventions to improve school performance, e.g. the cumulative expenditure of educational establishments was far lower than that of **PISA-D countries**: in 2013, **Senegal** was at US $ 6,818 equivalent, **Ecuador** was at 14,011. and **Paraguay** at 13,756. The largest expenditures are found in **Singapore** (130,611) and **Luxembourg** (187,459); The best performing countries in PISA-D have invested much more in education than others, e.g. **Ecuador** and **Singapore;** The contribution of parents and parent organizations remains as low as all **PISA-D countries** except **Cambodia.**

Latin American countries

Ecuador	**Low student performance**	to *involve parents in school activities*	**SUMMA (2018); Avvisati et al. (2014); Berlinski et al. (2016); Dizon-Ross (2018):** Greater participation of parents in school activities has positive effects on student performance e.g. **U.S., Canada, UK, Germany, Chile, Peru** and **Mexico**.
		to *encourage tutoring or peer support*	**Gingsbury-Block (2006); Leung (2014); Washington State Institute for Public Policy (2017):** Peer tutoring has positive impact on student performance in countries, such as **Brazil** (SUMMA, 2018a), **U.S.**, (Bernstein, 2009) and **UK** (Maxwell,2014).
	Gender gap in math performance	to *encourage girls' learning in math*	**What Works Clearinghouse (2018):** Academic skills of students, esp. girls with lower performance are expandable and improvable. **Halpern (2007):** Teachers are recommended to understand and communicate to their students that math and science skills can be improved through constant effort and learning. **Dweck (2006); Doolittle (1989):** Students who have a good performance but who see their cognitive skills as fixed are more likely to experience greater discouragement, lower performance and reduction of their efforts when encountering difficulties. **Blackwell (2007):** Stereotypes about girls' innate abilities are usually more common in the context of mathematics. **Good (2003):** Students who tend to view their skills as expandable try to increase performance despite encountering difficulties.
	Investment in education	to *focus on the distribution and use of educational resources* *(1) focus on educational areas and institutions with limited resources.* *(2) compare policies & practices in educational spending.* *(3) track & assign students based on level of learning.* *(4) provide complementary inputs e.g. budgets for materials & class plans.*	**Burtless (1996); Nannyonjo (2007); Nicoletti and Rabe (2012); OECD (2013, 2016a); Suryadarma (2012); Wei et al. (2011):** Once a high level of resources has been reached, more resources do not necessarily imply better learning outcomes. **Bressoux et al. (2009); Lavy (2012); Henry et al. (2010); Schenzenbach (2007); Willms (2006)**: Increasing the educational resources available to students and colleges with a low socio-economic index is beneficial at the level of student performance, and also compensates for inequalities in education. **OECD (2016a):** in **the highest performing educational systems**, resources are usually distributed more equitably among schools with a high and low socioeconomic index. **The best educational systems** seek ways to reallocate resources to address the circumstances of students who lack the material and human resources. **The Learning Better (IDB, 2017)** indicates that tracking is an effective resource reallocation practice. **Busso et al. (2017)** indicate that complementary inputs should be provided with sufficient support and monitoring to give good results.
	Learning differences	to *focus on teachers and professional development* *(1) attract, develop & retain high quality teachers.* *(2) provide teacher development opportunities;*	**Busso et al. (2017); Araujo et al. (2016); Hanushek and Rivkin (2012); Learn Better (2017); Vegas et al. (2016); Vegas and Ganimian (2013):** Teachers and their interactions with students play a key role in the learning process. **Vegas et al. (2016); Vegas and Ganimian (2013); Learn Better (2017):** policies related to teacher selection, training, support and motivation are fundamental, of which the training programmes at work and the use of monetary incentives have proven to be effective. **Busso et al. (2017):** teacher training programmes may be more effective in contexts where teacher preparation is weaker; there is also evidence of positive impacts of interventions that seek to develop the planning, monitoring and supervision skills of school management.
Guatemala	**School drop-out**	to *establish strategies to support students' educational trajectory* *(1) ensure the transition between levels.* *(2) develop educational support programmes e.g. food, school supplies & educational bag;* *(3) review the grade repetition policy.* *(4) retain students at risk of dropping out.*	
	Grade repetition	to *review grade repetition policy* *(1) take into account transition processes & improvement strategies.* *(2) commit training & teaching required to work with students who lag behind.*	**(Jimerson, 2011):** repeating one or several grades can be costly, both for the State and for families, but mainly for the student because of the negative effects on academic performance.

		*to **implement alternative models of basic education***	
		(1) develop a curriculum based on learning needs.	
		(2) train teachers who can develop a flexible curriculum.	
		(3) provide relevant material & didactic resources.	
	Academic performance	*to **maintain performance in reading & make improvements in math & science*** *(1) promote early intervention.* *(2) develop skills as measured by the PISA test items.* *(3) perform formative evaluation.* *(4) monitor and evaluate the implementation of the National Base Curriculum (CNB).* *(5) review the elementary curriculum.* *(6) support for students from private school.*	**Hat-tie (2007)** shows that formative evaluation, esp. the feedback, is essential to prevent non-learning.
	Health and well-being	*to **increase health and well-being*** *(1) strengthen health throughout child's development.* *(2) incorporate health & well-being learning into the CNB.* *(3) promote emotional & physical wellness programmes.* *(4) evaluate projects related to health & well-being.* *(5) expand the school feeding law to the secondary level.* *(6) ensure the health & well-being of teachers.*	**The PISA-D results** show that comprehensive and transversal attention to the welfare of the student population is neceassry.
	Student commitment	*to **increase the percentage of students at level 2 or above by 10%*** *(1) develop student life plans.* *(2) perform peer work.* *(3) link the results of national assessments to improvement programmes.*	
	Low investment in education	*to **increase resources to reach 7% of GDP*** *(1) make decisions based on data.* *(2) priortise the expenses.* *(3) increase the education budget.* *(4) create a policy of educational resources.* *(5) establish minimum criteria of spending for a decent school.*	The accumulated expenditure per student of Guatemala is less than 10% of the average accumulated expenditure of the **OECD countries** and is less than a quarter of the average accumulated expenditure by students from **Latin American countries** that participated in **PISA 2015** and **PISA-D**. The **PISA-D data** show that the official rural establishments have the most needs and problems, so they do not offer spaces worthy of learning.
	Low parental involvement	*to **increase family and community support*** *(1) involve parents in the classrooms.* *(2) create family support programmes.* *(3) strengthen family and community support in reading.* *(4) monitor children's learning using technological device.*	
	Quality education	*to **increase quality education*** *(1) strengthen in-service teacher training programmes.* *(2) extend the accompaniment processes to improve teaching.*	

Paraguay		*to **increase learning time*** *(1) ensure the optimal use of learning time.* *(2) develop guides for teachers in classroom teaching.*	**PISA-D results** show that hours of instruction for 14-year-old students in Guatemala is similar to the hours established by **Peru**; Guatemala has more accumulated instructional hours than **Finland** but less than **Chile**: the results of students do not correspond to the time due to the lack of quality instruction. **PISA-D results** show that absenteeism, unpunctuality and interrupting the development of a class have negative effects on learning.
	Inclusive environments	*to **increase learning environments*** *(1) implement national citizenship strategy.* *(2) prevent sextual harassment & other forms of violence.* *(3) establish alliances & promote tolerance towards diversity.*	**SDG 4.7:** ensure all learners 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 sustainable development.
	Scarcity of resources vs. Immediacy of improving competencies	***Short-term policies*** *(do not cost much and can generate change)*	
	Loss of learning time	*to **make effective use of time*** *(1) ensure attendance & punctuality of students & teachers.* *(2) improve the quality of classroom activities;*	**Alfaro et al. (2015)** indicate that the amount of time is only relevant if it is used to involve students in academic learning aligned with their needs. Data from **national and international assessments** shows the harmful effect of absenteeism of students in students' academic performance. Data from national and international assessments show the harmful effect of absenteeism on student performance.
	Violence and insecurity	*to **mitigate the intensity of violence*** *(1) combine guidelines and information and awareness campaign.*	
	Insufficient family & community support	*to **strengthen family and community support*** *(1) inform the benefits of parental involvement.* *(2) identify barriers and improve communications.* *(3) develop interactive activities e.g. joint reading.*	**Avvisati et al. (2014); Berlinski et al. (2016); Cerdan-Infantes and Filmer (2015):** Schools can improve communication with families if the communication channels that are best known are adopted and use a clear and inclusive languages in their communications. **Bauer et al. (2018)** shows that communication with parents about student attendance is effective. **Kim (2007; Kim and White, 2008; Sénéchal, 2006; Sénéchal and LeFevre, 2002, 2014):** The involvement of parents in the exercise of reading is more significant than the simple availability of texts to students, especially in childhood.
	Grade repetition	*to **correct the perception of grade repetition*** *(1) use awareness campaigns.* *(2) review the repetition reduction policy.* *(3) implement remedial strategies.* *(4) design and implement a structured and well-detailed curricular programme;* *(5) conduct regular formative and systematic evaluation*	**Brophy (2006)** shows the harmful effects of repetition on learning, student retention, and such dimensions as self-esteem, attitudes towards school, interpersonal relationships with classmates, absenteeism, and behavourial problems, as well as on the efficiency of public spending on education. **MEC (2013):** «Teaching Sequences» offers standardised instruments that allow the identification of each student's progress according to performance levels and areas of competence.
		Medium- and long-term policies	
	The learning problem	*to **improve the coverage and retention*** *(1) improve early childhood education.* *(2) improve the quality of education.* *(3) adopt a competency-based approach.*	The literature shows that both good bad learning opportunities have cumulative effects in the long term. Children who live in a context that contributes to their learning receive stimuli that are up to their abilities and are consistently accompanied by an adult during their growth, reach higher levels of development **(Willms 2015)**. **UNICEF (2017)** shows that a person's development opportunities start at the same conception, with prenatal care inherent in the period of pregnancy and become crucial in the first 1000 days of life.
	Not enough spending	*to **increase investment in education*** *(1) invest in all levels of education.* *(2) review teacher salaries and benefits;*	**PISA-D data** show that teachers themselves declare that they are not satisfied with the remuneration received. **Bruns and Luque (2015)** show that the link between remuneration and professional development is strong. There is very narrow space for teacher professional development within the ministry (**MEC, 2018**).

	Quality of education	*to **optimise the use of material and human resources*** *(1) conduct cost analysis and guarantee mechanisms.* *(2) review and improve teacher training systems, including access requirements.* *(3) accredit teaching training career.*	**Burns and Luque (2015):** The review and improvement of teacher training systems, including access requirements and the levels of demand to access teaching positions, together with the improvement of work incentives, would contribute to the hierarchy of the teaching career, attracting the best students and professionals.
	Inequity	*to **develop and establish standards of educational quality*** *(1) include all areas: learning, professional performance, school management, infrastructure & resources.* *(2) involve all actors: state, governorates, municipalities, communities, families, teachers & students.*	

Overall, inconsistencies and contradictions are prevalent both across and within the national reports. For example, there is a notable divergence of policy recommendations across the countries, which suggests different approaches to and some domestic influence on the composition and selection of policies. But, since the reports were largely based on inputs and guidance from the OECD and its private contractors, they also share some common parameters in terms of the framing of policy actions. The first parameter is the key elements identified within the EPA, especially those focusing on establishing the foundations. The Guatemala report can be seen as an example of adhering tightly to the EPA narrative in its framing of policy actions. The second parameter is the tendency to categorise policy actions in accordance with short-, medium- and long-term effects on education systems. This tendency is prevalent in the Cambodia, Zambia and Paraguay reports. In particular, the Cambodia report refers to policies that do not cost much but can have impact as 'low-hanging fruits' and 'quick wins' (2018, 135–137). Similarly, the Zambian report also identifies some 'new policies' that do not cost much and can have a positive impact, along with 'long-term policies' that can have impact in the long run (2018, 126–127).

Below, I examine in more detail the aforementioned inconsistencies and contradictions, with particular attention to the forms of policy actions recommended and the various sources of evidence cited to support those policies. Afterwards, I provide a critical reflection on the application of the EPA across the national reports.

Firstly, there was a similar effort in addressing grade repetition and student drop-out in the Cambodia, Zambia, Guatemala and Paraguay reports and this echoes the main report; however, the effort was not reflected in the Senegal report, where the grade repetition rate (50%) is the highest among PISA-D countries. In terms of the sources of evidence cited to support this policy, the Cambodia and Zambia reports rehearsed the OECD narrative on the effect of grade repetition in delaying students' entry into the labour market, whereas the Guatemala and Paraguay reports drew evidence from scholarly research and paid more attention to the effect of grade repetition on students' learning attitudes and behaviour. To some extent, the different sources of evidence explain the divergence of policy actions, but, in many cases, they conflict with the OECD narrative or those presented in other reports.

For instance, the main OECD report suggested that 'gender differences in attainment and proficiency are not large and are sometimes in favour of girls across PISA-D

countries' (2018c, 16). Obviously, this was not the case in Ecuador, where girls' performance in math is significantly lower than the performance of boys. This gender gap, which translates into one year of schooling, is identified as a major problem in the country. The Ecuador report thus noted that the difference indicates not only the necessity for girls to develop their math skills, but also the criteria for ensuring equality and equity in the teaching and learning process. This call for improvement is in contrast to the OECD's retrospective view that PISA-D results reflect countries' efforts to implement gender equality policies over the past decade.

Similar problems can also be observed in the area of investment in education. The OECD has repeatedly stated that it is not how much resources are spent but how those resources are used that matters, and that the right education likely compensates for low investment in education (e.g. Schleicher 2018). In the Cambodia report, there was the acknowledgement of the scarcity of resources available in the country, which suggested the avoidance of policy actions that are costly or require further investment. In this regard, the report identified five 'low hanging fruits' to help rectify the current situation. These were considered as 'extra actions from schools [that] can bring about substantial changes in student learning at the least expense of finance' (MoEYS 2018, 135). Accordingly, apart from reducing grade repetition, the other most effective approach was to align learning activities with PISA test items. The major source of evidence cited to support this action is drawn from the OECD (2013a, 2014c) reports on the successful experience of Asian countries in PISA. The following statements provide an illustration of how this policy action was suggested:

> Clearly, learning opportunity is limited. More than that, there is also a relatively large amount of instructional time loss. Having a longer school day is a response to this challenge, but requires more investment, which is a challenge for a country investing less than 3% of its GDP in education. Creating extra learning opportunities through increasing learning tasks per lesson/chapter and align them with the competency skills is an alternative solution. This can be done in the form of developing workbooks, extra creative tasks for students to practice in the class and at home, and the resource banks for teachers and students from the early grades. *All need to be aligned with the competency-based curriculum or PISA test items.* The *successful experience of Asian countries* such as South Korea, Hong Kong, China Macao, *Vietnam* and Taiwan through complementary education *indicates that this can be done*, particularly with a strong engagement of parents and the community in the learning process of the students both at school and home (OECD, 2013; 2014). (MoEYS 2018, 138, my emphasis)

Clearly the success of those Asian countries in PISA is not the direct result of aligning students' learning with PISA test items. In Chapter 5, one of my interviewees has already revealed that Vietnam performed really well on PISA primarily because they tested the

elite groups – only those who can afford to go to school. In the Cambodia report, right above this alignment-with-PISA policy, there was also the recommendation to make learning a serious goal and establish a strong learning culture (as 'quick wins'). The evidence was again drawn from major East Asian countries, such as Beijing–Shanghai–Jiangsu–Guangdong (China), Macao (China), Hong Kong (China), Chinese Taipei, South Korea and Japan – top performers in science in PISA 2015:

> The deep-rooted learning culture, perhaps derived from Confucius school of thoughts, plays a key role in their success. Many argue that resources are key to their success. But evidence shows that it is not about how much resources are spent but about how those resources are used (OECD, 2013; 2016a; 2016b). U.S., UK, Australia and the other OECD countries, for example, spend significantly more but lag behind South Korea, Finland and even New Zealand in reading performance. Building a strong learning culture, in fact, makes a difference. (ibid., 137)

Here, beyond even the fact that Confucian thinking is not readily transferable to a country, where 85% of those population believe in Buddhism, the challenge lies in using evidence drawn from countries that do not share similar socio-economic conditions with Cambodia (i.e. East Asian countries, the U.S., the UK, Australia and other OECD countries). These more affluent countries have already established the 'foundations for success', whereas in Cambodia, schools are not well-resourced, reflecting the lack of foundations. Thus, further investment in education is self-evidently necessary in order to allow policy actions to take effect. From this, the selective use of evidence (i.e. examples of 'best practices') merely serves to legitimate preferred policies that enhance the OECD's governance in education by encouraging alignment of learning activities with PISA test items.

The Cambodia report goes even further, considering the alignment as a long-term goal to be integrated into early childhood and basic education. In this regard, the logic of the EPA seems to be used to justify early interventions. As the report states, 'PISA-D results suggest evidence that poor basic skills […] in the lower grades is what truly matters to the overall student performance in secondary education since PISA-D measures the cognitive skills students develop across ages' (ibid., 140). Thus, while educational investment in the country is low, 'increasing instructional hours or expanding school days to promote practices of cognitive skills for the low grades' is regarded as a long-term solution. 'To achieve this goal, the ministry needs to focus on incorporating a concept- or a competency-based teaching and learning modality into the textbook and aligning them with classroom assessment' (ibid.). This practice is planned to 'support the

implementation of the new curriculum and develop competencies for Cambodian students to prepare better for international student assessments such as PISA' (ibid.).

Regarding the evidence for early interventions and for teaching to the test, the Cambodia report drew not on evidence from PISA-D, but on the OECD projection study conducted by Hanushek and Woessmann:

> [t]he investment in early childhood and basic education (K–9) plays a key role in achieving universal basic skills and in maintaining economic growth. The projection study by OECD 2015 shows that increasing average achievement in current students by 25 PISA score-points has a uniform effect on all countries' GDP by 30% over the next 80 years if there is a 100% enrolment (Hanushek and Woessmann, 2015). (ibid.)

In Guatemala, there was a similar emphasis on developing the skills measured by PISA. The national report suggested that this intervention better starts in the early grades to establish the foundations in students. However, the role of assessments in this respect was considered as essential to prevent non-learning and to reduce gaps between students due to differences in sex, geographical area or socioeconomic level, rather than based on the OECD's projection.

In Zambia, the national report concluded that the benefit for the country's participation in PISA-D is not clear, except that students performed poorly in all of the three subjects. It stated that 'the persistent problem of low learning achievement levels by Zambian learners has been empirically documented since the late 1990s' (2018, 128). 'Subsequent surveys have repeatedly pointed to the problem of low learner achievement over the years' (ibid.). At the regional level, the SACMEQ (Southern and Eastern African Consortium for Monitoring Educational Quality) results 'have equally confirmed the persistence of low learning achievement levels by learners in Zambian schools, to the extent that the country compared poorly even in the sub-region' (ibid.). As such, the major suggestion that the report made was to increase the country's political will to address critical and specific recommendations that have been repeatedly reported in the various studies, including SACMEQ and, most recently, PISA-D.

From intertextual readings, Zambia's experience with PISA-D confirms the critiques raised by the UNESCO representative about Cambodia endorsement of PISA-D. I provide the extract below:

> Despite knowing the fact that, for years, the education system continues to struggle implementing the instructional hours stipulated in the curriculum, teaching and learning is being delivered by underqualified and under-supported teachers in crowded classroom settings, with shortages of textbooks and learning materials, this is going forward. Further, the results from existing learning

assessments, examinations etc. are already pointing to the fact that massive efforts and investments are needed in improving reading and mathematics from early grades on. So why get distracted or distract? I would rather use the scarce money to provide…teaching guides to the teachers. (UNESCO representative, interview conducted March 2017, cited in Auld, Morris and Rappleye 2019, 10)

Secondly, although there has been a marked difference across the reports on the sources of evidence cited to support policy actions, most of these sources are still based on previously published materials and references to practices that are deemed to work in other countries. In particular, the Cambodia report relies heavily on OECD reports and references to countries which perform well on the main PISA, whereas the Ecuador and Paraguay reports draw on a wider range of research articles, many from regional sources and evidence from national assessments. However, by and large, there was notably limited reference to evidence drawn from other PISA-D countries. This tendency to avoid comparison between PISA-D countries, combined with the prevalence of references to high-performing or fast-improving PISA participating countries, undermines the OECD claim that 'the results of PISA-D allow participating countries to determine whether their policies differ from those of countries with a *similar* social and economic context' (2018c, 11, my emphasis).

Furthermore, given the reliance on high performers on PISA and pre-existing research as the sources of evidence, the policy recommendations were also loosely connected to the actual PISA-D data. This further challenges the OECD claim that 'assessment results provide policy makers with data and evidence that can be used to determine what they can do to improve their education systems and, ultimately ensure that their students acquire the skills needed to succeed in school and in life' (ibid., 4). These controversies confirm my interpretation of the origins of PISA-D and the OECD's global education ambition (shown in Chapter 4). In other words, the actual outcomes of PISA-D become more meaningful when understood from the OECD's point of view. For the OECD (2018c, 3), PISA-D has been successful, because it has helped the OECD to increase the 'resolution' and 'relevance' of the PISA instruments so that they can be tailored to any other low performers of the global PISA community. 'Resolution' can be seen as relating to adding the lower rungs on the PISA ladder and 'relevance' in terms of testifying the application of the EPA in low-resourced settings.

Another related benefit is that PISA-D has helped the OECD to include more 'diversity in policies and practices, enriching analyses by a greater range of points of comparisons and also increasing the opportunities for peer learning' (ibid.). The aspect

of diversity and inclusion are illustrative in the PISA-D national reports, in particular, the divergence of policy options and the greater range of reference points of comparisons. However, whether or not those options and comparisons are meaningful and context-relevant is another matter, at least for the OECD. The primary goal that PISA-D serves to achieve is to transform and upgrade OECD policy instruments and to ensure their accessibility and 'relevance' to low- and middle-income countries.

Considering the lack of data infrastructure particularly in low-income countries, these two aspects are essential for overcoming the barriers to countries' participation in PISA: the lack of capacity to implement the assessment and a lack of experience in using PISA data and results. Now with the data infrastructure in place, the OECD is better able to place all countries on one common development scale and modulate them according to PISA metrics. PISA-D, therefore, is a self-serving device, rather than serving the needs of low- and middle-income countries. This was especially evident in the promotion of policies designed to reduce grade repetition and assessment capacity; if pursued such policies would primarily serve to weaken the barriers which hinder the implementation of PISA testing.

Thirdly, while the lead analysts adopted the EPA as the guiding framework to develop their national reports, it is readily apparent that the framework has not been applied with a reflexive understanding of the implicated assumptions and values. The process entails an uncritical appropriation of a developmental model that is imposed by the OECD and its private contractors and is found on 'planet North'. To draw on Nóvoa (2018, 551), the generalisation of global solutions – as a pragmatic approach – is 'based on the false idea of consensus on the aims of education and the paths to achieving them'. However, in agreeing with Nóvoa, this consensus is also made possible because policy makers and the lead analysts in PISA-D countries have accepted 'an empirical status' with regard to their own studies. As such, 'the theoretical and epistemological status was left to the "North"' (ibid., 557) and the 'empirical status' was merely used to consolidate the academic rigour of the former.

In the case of the EPA, Willms has since been utilising PISA-D countries' experience as "cases" or "exemplars" to demonstrate the utility of his model in the low-resourced settings. For example, Willms presented the EPA at the Conference on 'Learning at the Bottom of the Pyramid' (held in March 2017) at the University of Pennsylvania. Later in Chapter 5 of the published volume *Learning at the Bottom of the Pyramid: Science,*

Measurement and Policy in Low-Income Countries (Wagner, Wolf and Boruch 2018), Willms promoted the EPA as the assessment strategy for supporting student learning in low-income countries. To this end, he referred to the adoption of the EPA by OECD and PISA-D countries as evidence to suggest its wider applicability:

> The approach is being used in a number of middle- and high-income countries, and has been embraced by the OECD and the countries participating in PISA for Development. Its use has implications for how large-scale assessments can be used to support student learning in low- and middle-income countries at the classroom and school levels, and how large-scale assessments can inform educational policy at the district or state levels. (2018a, 99)

Another notable publication is the recent UIS report entitled *Learning Divides: Using Data to Inform Educational Policy*, authored by Willms (2018b). In this report, Willms again elaborates on the EPA, highlighting how it can be applied to help monitor SDG 4 targets and to discern effective interventions. The content of the report draws largely on that developed for the PISA-D contextual questionnaires (e.g. Willms and Tramonte 2015). In particular, I notice that in this report, when introducing the key factors that impact equality and equity, Willms stated that they 'were identified by the countries participating in PISA for Development to be the most important in their contexts' (2018b, 17). However, this is not true, because most of the factors, such as gender, immigrant status, language, and socioeconomic status, were already included in at least four cycles of PISA (see e.g. PISA 2015 context questionnaires framework, OECD 2017c), and, so, the factors (as core indicators) were merely modified to relate to the contexts of PISA-D countries, rather than identified by the countries themselves.

Nevertheless, since now PISA-D is 'so successful' (as Schleicher celebrated at the latest PISA-D International Seminar), all the instruments that have been developed have become meaningful and relevant. And for Willms, an 'empirical status' relating to PISA-D is enough to capitalise on.

Summary

The chapter examined the outcomes of PISA-D, including the assessment instruments, the findings from the school-based assessments, and policy recommendations. PISA-D aimed to overcome longstanding challenges which prevent low- and middle-income countries from engaging in PISA. It has been successful, from the OECD's perspective, for three reasons. First, by adding new proficiency levels (e.g. 1a, 1b and 1c) to the lower

end of the PISA scale and anchoring PISA Level 2 to the basic minimum proficiency identified under SDG 4.1.1(c), the OECD placed these countries on the PISA development scale. This outcome helps to prepare for universal participation in PISA. Second, through incorporating the EPA to adjust the contextual questionnaires and to frame and interpret the findings generated from these assessment instruments, the OECD established the reference points (or variables, such as resources, quality instruction etc.) for international comparisons. This outcome also helps the OECD to identify and suggest effective policies. And third, by prioritising interventions which encourage alignment with PISA test items, the OECD ensures their involvement in future PISA cycles. The third aspect is also essential in the sense that such an alignment at early grades have the potential for enlarging the PISA 'configuration' (OECD Interview#2, 2018) by opening up the vertical, longitudinal dimension of the global assessment regime.

While PISA-D was positioned as a pilot project, its experimental success was preordained. The piloting status was used to validate and legitimate the extension of PISA into low- and middle countries (and thereby 'throughout the world'; OECD 2019d). PISA-D is a self-serving device, which addresses squarely the core elements of the OECD's development strategy for 2011 and beyond. Although PISA's legitimacy has largely derived from its portrayal as an objective and competent measure of education quality, my examination of the technical procedures and the assumptions implicated in PISA-D shows that the actual implementation was intrinsically problematic and involved undertakings that 'fabricate and manage' (Grek 2013, 699).

Certainly, this is not empty theatre, with the outcomes assembled and disseminated through working papers, national reports and public seminars, bringing together OECD officials and mainly those engaged in the project who vouched for its success. The reality is, although many of the challenges had been acknowledged, they were not resolved. So, for example, there were no new test items developed because that would interrupt the comparability between PISA-D out-of-school, PISA-D in-school and the main PISA tests. This is a bottleneck and a good illustration of how the expansion of PISA has limitations. The application of the EPA was also symbolic and even self-contradictory, as it operates on a different logic and is unsuitable for under-resourced settings, especially out-of-school settings.

Moreover, as already discussed extensively, policy lessons derived from the findings of PISA-D are uncritical and superficial, confirming what Nóvoa (2018) terms the

'solutionist drift' in comparative education. The process and outcomes of 'peer-learning' also echo the depiction of PISA as a form of 'banal imperialism' (Silova and Auld 2019) which places PISA-D countries remote in space and time and, so, the enhanced instruments and emulating the policies of high performers (or fast improvers) on PISA would provide the 'only way up'. All of these undermine the OECD's claim that PISA-D is truly successful, except for its own interest. The complexity of the issues and challenges that have emerged within and across PISA-D participating countries defy the OECD's tidy boxes. As noted in Chapter 4, the OECD was aware of the limits to its approach to education and development and, therefore, set out the strategy to prioritise a coherent approach by drawing on multi-disciplinary expertise and a governance toolkit. However, it has failed, as we observed from the experiments of PISA-D: the facts on the ground have continually challenged the OECD's definitions (e.g. of 15-year-olds or 'out-of-school' youth) and standard operating procedures, which form the foundations of its expertise and authority.

Chapter 7. OECD Governance, Organisational Legitimacy and Banal Imperialism

> Empire is materializing before our very eyes.
>
> Empire establishes no territorial center of power and does not rely on fixed boundaries or barriers. It is a *decentered* and *deterritorializing* apparatus of rule that progressively incorporates the entire global realm within its open, expanding frontiers. Empire manages hybrid identities, flexible hierarchies, and plural exchanges through modulating networks of command. The distinct national colors of the imperialist map of the world have merged and blended in the imperial global rainbow.
>
> —Michael Hardt and Antonio Negri, *Empire*

Introduction

This thesis recognised the significance of PISA-D and critically analysed its origins, the legitimation strategies which frame it, and its outcomes. Three major research questions have guided this study. What are the origins of PISA-D and how have they shaped its contemporary form? How has PISA-D been promoted and legitimated, and what strategies have been employed by the OECD to establish its role in monitoring SDG 4? How has PISA-D addressed the challenges that low- and middle-income countries face in engaging in ILSAs, how are the assessment results generated and compared, and what policy insights does it provide for the piloting countries?

Through documentary analysis, semi-structured interviews, and secondary analysis of the literature, this thesis has demonstrated that the OECD's story is one of successful expansion, but that the expansion has been achieved without addressing the fundamental and longstanding challenges which have faced the use of ILSAs in low- and middle-income countries. Legitimacy management has played an essential role for the OECD as it expanded its role and developed PISA-D. However, within these processes of expansion and legitimation, the PISA-D countries are construed as backward and needing to learn from other countries participating in PISA, especially those deemed to be high performing or fast improving. The consequence is that the promise that PISA-D would be 'contextualised' for low- and middle-income countries was not realised.

In this concluding chapter, I will revisit the previous chapters, showing how they are connected and how they link to the wider literature, with a focus on the themes of global educational governance, organisational legitimacy, and banal imperialism. I will also reflect on the OECD's recent educational activities in response to the global COVID-19

pandemic and how these activities continue to exclude out-of-school youth. Finally, I will highlight some limitations of this study and propose areas for future research.

Summary of chapters

The introductory chapter began with the background of the global education agenda being shifted after 2015 to emphasise the quality of education and identify minimum standards of quality for all countries to be delivered by 2030. This chapter then mapped out the strategic work of the OECD as it expands PISA into low- and middle-income countries and promotes it as a global learning metric. PISA-D is essential for the OECD's engagement with these countries and with the SDGs. This chapter highlighted the significance of PISA-D and explained how it differs from the OECD's other assessment programmes. The rest of the introductory chapter elaborated on the major research questions, their potential contributions to the literature, and the structure of the thesis.

Chapter 2 situated PISA-D within three relevant bodies of scholarship to understand its distinctive features. The chapter firstly drew on insights from CE literature to understand PISA-D as a form of 'applied' CE. This specific form, in contrast to the academic traditions of CE, is based on shallow philosophical and methodological grounds. Therefore, I suggested that PISA-D, which aims to compare all types of countries with vastly different conditions and needs, is problematic and oriented towards consolidating the status quo, including its injustices and inequalities. This part of the review provided a further basis for understanding the policy actions promoted in the PISA-D reports (examined in Chapter 6). Moreover, the review referred to the two major theoretical approaches in the CE literature, i.e. world culture theory vs. local agency/meanings, which help to explain the rise of the ILSA phenomenon. However, I suggested that these approaches are inadequate for understanding the global expansion of PISA, as their focuses of attention are not on the powerful role of the OECD – as an agent in the field of global governance and as a producer of a global testing culture.

This thesis recognises the powerful influence of the OECD. Therefore, the second part of the literature review (LR) incorporated perspectives from IR scholarship, namely realism, neoliberalism and constructivism. I argued that the sociologically informed constructivist approach, which perceives IOs as bureaucracies and legitimacy managers, has been particularly helpful for making sense of the OECD's increasingly prominent role

in global educational governance and, therefore, the contemporary form of PISA-D (shown in Chapter 4). The constructivist approach provided insights into the major sources of power of IOs, i.e. rational-legal authority, delegated tasks, moral claims, and expertise. It also shed light on the changing dynamics of relations between the OECD and other prominent agencies governing global education, e.g. UNESCO and UNICEF. These insights were helpful for identifying and interpreting legitimation strategies employed by the OECD to promote PISA-D and to establish a clear role for itself in monitoring SDG 4 (presented in Chapter 5).

Given that the OECD's major claims in promoting PISA-D are centred around improving economic growth and educational quality, the third part of Chapter 2 delved into development studies to understand the relations between PISA and these aspects of development. The review first looked at studies which examined the impact of PISA on future economic growth and then referred to those examining the positive impact of PISA on student learning. However, this part of the review generated mixed results with regard to claims that improvements in PISA scores were a direct contributor to development. More crucially, it also suggested that the OECD has an imperialist tendency in its attempts to enrol low- and middle-income countries into one global assessment regime and to impose its own unequivocal logic, albeit this logic is operationalised through technical and economistic approaches and is disguised as education performance data. Chapter 2 did not attempt to validate any one theoretical strand and instead, tried to refrain from a single reading by acknowledging the multiple perspectives on reality and problematising the OECD's normalised versions.

In Chapter 3, I described the research strategy and design by adopting a critical hermeneutic approach. I developed the research design based on the belief that multiple, intersubjectively constructed truths about social reality exist. Therefore, I perceived and analysed the OECD and PISA-D through interactions with key actors and organisational documents, as well as the literature. I considered my positionality as being integral to the meaning-making processes; my understanding of the OECD and PISA-D developed and deepened, following the path of a hermeneutic circle–spiral, during the course of data interpretation and analysis. The critical part of my analysis lies in the effort to unmask the dominant ideological structures underpinning PISA and PISA-D and how they serve to expand the OECD's governance potential and consolidate the existing injustices and inequalities. Accordingly, I applied this critical hermeneutic approach to reveal how the

OECD acts as an empire and uses an imperialist discourse to construct the PISA-D countries as backward and needing to learn to improve themselves through the path dictated by PISA (as shown in Chapters 4, 5 and 6).

Chapter 4 traced the origins of PISA-D through a periodisation of the OECD's changing agendas and approaches to education and development. Three major periods were identified: from the 1950s to the 1990s; from the late 1990s to 2010; and from 2011 to 2030. This chapter devoted considerable space to the latter period to highlight the OECD's emerging governance approach which emphasises education for sustainable development. This approach, in which PISA shifts to become part of a 'configuration' (as described by the OECD staff member interviewed), significantly transforms the previous one-size-fits-all approach in which PISA served as the universal standardised instrument. I suggested that this emerging approach is more inclusive and flexible and potentially more digital and individual, as the 'configuration' incorporates OECD's multidisciplinary expertise and a governance toolkit which can be tailored to each individual country regardless of geography and levels of development. However, I argued that this 'configuration' is premised on universal similitude and expropriates others' cultures, beliefs and knowledge. The OECD modulates it through a single hierarchical command but places low-income countries as remote in space and time. In this chapter, I also suggested that while the OECD's story is one of successful expansion, albeit with limits and challenges. Legitimacy management was thus essential for securing organisational survival and expansion.

Chapter 5 focused on identifying the legitimation strategies the OECD employed to promote PISA-D and to establish its role in monitoring SDG 4. I identified six critical strategies: (1) *building affiliations*; (2) *shaping and (re)framing SDG 4*; (3) *establishing credibility*; (4) *creating demand*; (5) *membership and belonging*; and (6) *persuasion in motion*. The starting point of my analysis was the OECD's intention to gain socio-political acceptability to help define and then measure SDG 4. I showed that, to this end, the OECD first highlighted the role of, and affiliated with, the specialised UN agencies (in particular, UNESCO and UNICEF) that are mandated to carry out these tasks. The OECD capitalised on its technical expertise in PISA and related policy analysis to contribute to the framing and monitoring of the education SDG. Having gained socio-political acceptability, the OECD then succeeded in establishing credibility for PISA-D by highlighting the collaborative relations established with the mandated agencies, the extensive coverage of

PISA, and the validity and reliability of PISA instruments. Moreover, the OECD used certain exemplars, e.g. Brazil, Vietnam and Peru, to help demonstrate how PISA can drive development in low- and middle-income countries and thereby serve as a global learning metric.

Thereafter, the OECD articulated a demand-led story, claiming that PISA-D responded to the demand from countries wanting to engage with PISA and to the need of the international community for a global learning metric to track progress on SDG 4. The OECD supported this story by elaborating the privileges and benefits that low-income countries receive if they too become members of the global PISA community. The suggested privileges and benefits include, for example, international comparisons, world-class testing expertise, peer-to-peer learning opportunities, and demonstration of public accountability. Finally, the OECD discourses continued to circulate at each PISA-D meeting and workshop and gained momentum at the PISA-D International Seminar in 2019 where the project was declared a success. I thus suggested that these meetings, workshops, and the seminar were key relational sites for the OECD to promote PISA and PISA-D around the globe – which I called PISA 'boosterism'.

While the major focus of Chapter 6 was to examine the outcomes of PISA-D, it also identified an additional OECD legitimation strategy: *piloting PISA-D for success* (see also Auld, Li and Morris 2020). The OECD had positioned PISA-D as a 'pilot project' to help avoid criticism by placing it within a process of investigative science. This strategy worked well with the previous strategies (identified in Chapter 5) as the OECD's repertoire of approaches for legitimacy management. However, in Chapter 6, I showed that the experimental success of PISA-D was preordained. Although PISA-D did serve some experimental purposes by adding new proficiency levels at the lower end of the PISA scale and introduced the out-of-school component, my interrogation fundamentally challenged the OECD proclamation that PISA-D was truly successful.

In reality, PISA-D has failed to address the longstanding challenges (e.g. including a sufficient number of out-of-school youth) that low- and middle-income countries face in engaging in ILSAs. The newly added test items primarily aimed to place all countries on one common scale through which the piloting countries could begin their ascent; the EPA is a fabricated developmental model used to adjust the contextual questionnaires and to interpret the results from PISA-D school-based assessments. The policy lessons presented in the reports did not appear to be derived from the actual data but were largely framed

by the EPA with evidence drawn from high-performing (or fast-improving) PISA countries; the main recommended policy interventions encouraged piloting countries' involvement in future PISA cycles and alignment of teaching and learning activities with PISA test items.

All of these points suggest that PISA-D was not successful, except for promoting the OECD's objectives. I thus argued that the predefined success of PISA-D prepares for the acceptance of PISA as a global learning metric and, therefore, the OECD's ambition of ensuring assessment for all. This ambition was also evident in the PISA-D post-pilot arrangements (shown in Table 4.2, Chapter 4) in which there was no mention of (resolving) any of the challenges that PISA-D had actually encountered during the implementation phases (examined in Chapter 6), for this would have undermined its status as a success and the OECD's legitimacy and authority (discussed in Chapter 5). Instead, through these arrangements, the OECD has primarily focused on how to use the PISA-D instruments to compose a larger PISA programme, which can be optimally applied to all countries, especially low- and middle-income ones. Resonating with Hardt and Negri' (2001) characterisation of empire, quoted at the beginning of this chapter, I suggested that the OECD acts an imperial power progressively extending its influence and authority to the furthest corner of the globe.

Contributions to the literature

In the previous chapters, this thesis provided new interpretations of the OECD's contributions to global educational governance, highlighted the legitimation strategies of OECD as it expands its role, and made clear how the PISA-D countries are construed as backward and needing to learn from other, especially high-performing, countries participating in PISA. I elaborate on these major themes below. Finally, I provide observations on the OECD's recent educational activities in response to the global COVID-19 pandemic, which has permitted additional interpretations of PISA-D and the operating logic of the OECD in safeguarding legitimacy and organisational survival.

To start with, this thesis contributes to the literature on global educational governance by identifying the OECD's emerging outlook on governance as a *comprehensive* approach to education and development (as summarised in the 4th column of Table 4.1). This approach operates at both organisational and educational levels. As my interview

participant described, PISA is now part of a 'configuration' that includes the OECD's multidisciplinary expertise and a governance toolkit which can be tailored to the needs of all countries. In Chapter 4, I illustrated how this 'configuration' was concretised in the form of a 'larger' PISA programme through the post-pilot arrangements of PISA-D. Then in Chapters 5 and 6, I further showed how the OECD mixes and matches different countries, among which, however, the PISA-D countries are positioned as backward – and as needing to learn from high-performing (or fast-improving) PISA countries to 'take off'. I suggested that while the OECD's governance approach has become more inclusive and flexible, it functions through a notion of universal similitude as the unitary force which positions different cultures and identities along one common scale and links them all through objective data. In this way, through its global assessment regime, the OECD becomes the hub of a global policy network that supplies data, advice and transfer of 'best practices'.

Resonating with Hardt and Negri's (2001, xii) description of the empire that has no established 'territorial centre of power' and 'fixed boundaries or barriers', I suggest that the OECD is, through PISA-D, embarking on a similar passage towards becoming an imperial power exerting a single hierarchical command and bringing all within its sphere of operation and influence. The source of this influence is rooted in the triennial performance of an 'international' comparative spectacle exercised through the display of indicators and rankings, which, as Nóvoa and Yariv-Mashal (2003) note, expropriates values, beliefs and knowledge. Within that 'spectacle' the illusion of diversity and free will has been successfully fostered as the empire constructs a continuous, smooth space across which subjectivities glide without substantial resistance or conflict.

However, as I have shown through the genealogy of PISA-D, this comprehensive approach and its changing mode of operation are developing out of the circumstances in which the OECD's policy expertise and statistical authority are being challenged. PISA-D illustrates well this dilemma: PISA was criticised for not being relevant to the realities of low- and middle-income countries, and the OECD was also aware that it could not possibly stretch PISA 'in an unlimited way' (as stated by the OECD staff member). Therefore, PISA-D emerged through the progression of PISA as OECD sought to leverage its technical capacity by synthesising multidisciplinary expertise and by engaging with developing countries and the SDGs.

The experiment of PISA-D is critical for the OECD to demonstrate its commitment to adjusting or contextualising PISA by taking into account developing countries' conditions and needs. While in reality PISA-D has failed to address the longstanding challenges that these countries face, it is strategic in placing them on the PISA development scale which improves the OECD's capacity for monitoring SDG 4. Moreover, through the repertoire of strategies that appropriates and legitimates PISA *for* development, the OECD gains moral authority for providing PISA as the global learning metric. As Ward stated at the PISA-D International Seminar in 2019, although the pilot was carried out under assumptions that are not based on real science, 'in the context of SDG 4 (leave no one behind), there is no substitute for assessing the skills of the whole population'. Similarly, Schleicher (2013a) stated that the OECD has 'an obligation to *try* to make this [i.e. PISA-D] work for the sake of development and to support the enhancement of global skills levels' (my emphasis). These statements, along with the extracts in Chapter 5, show how the OECD has presented its work on PISA-D as socially valued activities and, therefore, itself as an organisation driven by moral goals.

This type of legitimacy relating to moral claims appears significant in the absence of clear measures of outcomes, in the sense that the pilot status of PISA-D is sufficient for demonstrating that the OECD is making a good-faith effort to achieve valued, albeit invisible, ends. As such, even though PISA-D has failed in its aims, in the end, OECD staff can celebrate that they at least tried, and the goals were worthwhile. While this finding is not novel, for example, other development agencies such as UNESCO and UNICEF have long asserted a humanitarian role, the OECD is distinctive in persuasive organisational communication and dissemination of its educational activities and their outcomes. As Pettersson and Popkewitz (2019) rightly observe, educational 'expertism' has changed over time, and today it is about communication rather than producing educational facts. Accordingly, in Chapter 5 I described the global PISA-D meetings, workshops, and seminar as embodying a specific form of marketing strategy – PISA boosterism – through which the OECD has widely disseminated PISA-D and mobilised test experts, policy makers, and even representatives from other agencies to help provide validation for its experiment.

Consequently, with the new instruments (i.e. low proficiency levels, the EPA-informed questionnaires, and the out-of-school component) carefully qualified and deemed successful, the OECD legitimates PISA-D as scientifically proven and proceeds

to compose a larger PISA programme that is well prepared to serve as the global learning metric. In these ways, PISA for development comes into fruition and the OECD is taking on a humanitarian role. Statistical authority and moral legitimacy appear to be held in close alignment. As discussed further below in light of the OECD's recent initiatives, this moral type of legitimacy helps the OECD to garner social support and allows its global assessment regime to continue to evolve and expand. The objective of the empire is to achieve 'assessment for all'.

By highlighting the legitimation strategies of the OECD as it expands its role and responds to both internal and external challenges, this thesis enriches the literature which holds an instrumental view of organisational legitimacy. Although organisational studies recognise the different forms of legitimacy, e.g. pragmatic, moral, and cognitive, that organisations seek to gain, maintain, or repair, few have applied these concepts to analyse the evolving role of the OECD in the field of global education policy (GEP).

As already noted, the GEP field in which the OECD operates is both normative and competitive. On the one hand, IOs as recognised authorities create social norms by investing meaning in information that coordinates values and prompts action. In the post-2015 era, this manifests in the normative understanding of a global 'learning crisis' that nurtures the necessity of identifying basic minimum standards of quality to measure and track learning (it is currently being amplified by the global pandemic). On the other hand, the 'thickening' of global governance institutions and technologies implies that organisational legitimacy has to be constantly managed and/or won.

For example, Edwards et al. (2018) have drawn on the literature on organisational legitimacy to interpret UNESCO's efforts to regain some of the legitimacy it lost in the decades before the EFA initiative. They conceptualise legitimacy as having three key components – sociopolitical acceptability, reputation and status – and argued that while an organisation can gain sociopolitical acceptability, it has no direct control over its reputation and status, as changes to them depend on an external social audience.

However, in the context of this thesis, my analysis of the legitimation strategies indicates the OECD's managerial role in manipulating public opinion. In other words, the OECD appears to have some direct control over its organisational legitimacy, including its reputation and status, to help achieve desirable outcomes. This can be clearly seen from the employment of strategies such as establishment of credibility for PISA-D, articulation of a demand-led story, and construction of a sense of membership and

belonging to the global PISA community, apart from the OECD's strategic use of media coverage to help reinforce the political impact of PISA (noted in the literature). Therefore, I suggest that it is through active manipulation, rather than conformity to expectations, that the OECD has successfully ascribed the meaning of PISA to development and strengthened its already prominent position.

The adoption of an instrumental view of organisational legitimacy provides insights into the modus operandi of the OECD in exerting power and influence. Moreover, it also helps to interpret the ways in which key agencies and actors (such as the OECD, the World Bank, UNESCO, and UNICEF) interact or compete with each other in the GEP field. For instance, Tikly (2017) describes the EFA agenda as a global regime of educational governance, which operates within a regime complex that is inhabited by powerful agencies and actors and is organised around a key set of principles, norms, rules and decision-making procedures. Within this complex, Tikly attributes the OECD's increasingly prominent role as a broker of education policy to the administration of the PISA international assessment regime, whereas UNESCO continues to draw its legitimacy from the human rights regime. Similarly, Ydesen (2019) observes that the contemporary governing complex in education leaves a substantial role for IOs with very different dispositions and instruments at their disposal.

My findings on the OECD's 'turn' to adopt a comprehensive approach to education and development, along with its efforts to gain moral responsibility for piloting PISA-D and monitoring SDG 4, suggests that it is moving from its human capital logic to embrace the rights-based regime. In the case of UNESCO, it has similarly kept up with the 'technical turn' (Elfert 2018) to build up its technical capacity in order to compete with other agencies, especially the OECD and the World Bank. These findings indicate that legitimacy dynamics remain critical for understanding change in organisations and the changing relations between them.

Having also taken an entry point from postcolonial scholarship, this thesis reveals how the OECD continues practices of domination and statistical 'othering' which extend beyond Western imperialism. In particular, this thesis recognises the 'banality' of these practices embodied in the routine production and dissemination of PISA and PISA-D. What makes these practices 'banal' is that they are increasingly taken for granted by the OECD, test specialists, and some policy makers, especially when presented in the form of seemingly neutral and objective data (Silova and Auld 2019). In the case of PISA-D,

these practices are similarly disguised and diffuse, and are further legitimised through the OECD strategies which emphasise peer learning and capacity building and mobilise resources to vouch for their moral authority.

Through the genealogy of PISA-D, I have revealed how this dominant reading is rooted in the OECD's mentality even as it evolved and as it continues to evolve. I have shown that while acknowledging their differences and idiosyncrasies, the OECD has ensured that the PISA-D countries tightly adhere to PISA technical standards and operating procedures to achieve international comparability. These have been primarily pursued through 'peer-learning' events and 'capacity building' activities, such as managerial meetings, workshops, and support visits, which have mixed and matched countries and translated participants into objects comprehensible through PISA's gaze (Addey and Gorur 2020).

However, the PISA-D countries – positioned at the lower end of the PISA ladder – are, by definition, remote in space and behind in time. Their identities in relation to progress, or being 'on track', can only be validated when they begin their ascent and demonstrate a certain measure of performance as judged by the OECD. This is the 'only way up'. While on their journey towards improvement, the 'low-performing' countries are accompanied by some 'high performers' (e.g. Japan and other East Asian countries), OECD averages, or 'fast improvers' (e.g. Brazil and Vietnam) as exemplars, the existence of an unequivocal order remains and is predefined by the OECD (one which is constantly shifting, as explained below).

Understanding PISA-D as a form of 'banal imperialism' opens up more fundamental issues relating to the value of engaging with PISA which have been well rehearsed in the literature. As this thesis shows, PISA-D reveals the same established problems inherent in PISA: as an 'applied' form of CE that is built on shallow theoretical and methodological grounds, and as a development project which hardly has to do with real development (or developments). From the constructivist perspective, PISA and PISA-D exemplify how the OECD applies rationalised versions of reality to facts on the ground. However, 'a complex world defies the bureaucracy's tidy boxes and neatly circumscribed division of labour' (Barnett and Finnemore 2004, 44). The OECD, too, has found that its definitions and standard operating procedures do not account for features of reality that threaten its ability to accomplish tasks. As such, it continues to evolve and to expand

areas of operation. Under the guises of objectivity and morality, however, we see that there is a clear imperialist ideology.

Recent development and response to the pandemic

Auld, Li and Morris (2020) observed that Schleicher has acknowledged two major limitations to the OECD's assessment regime. He first questioned the relevance of the DeSeCo framework, which underpins all the assessment programmes, for capturing the realities of the contemporary world:

> The current PISA is the reflection of the framework that we developed in 2000. Now we need a new one, and that will guide the development of PISA from 2024 onwards. *The current PISA doesn't include many elements that are important…* We need to do better on social skills, on creative skills, so the new framework – the 2030 [Learning] Framework is really about developing new ideas for what PISA should assess. PISA-D mostly […] developed easier items that are available for all sorts of contexts. Education 2030 is about what should PISA assess, and what are the skills that are really important for success. PISA-D is just making the existing PISA more adaptable, but it is not about developing new skills. (cited in Li and Auld 2020, 11, original emphasis)

He then questioned the capacity of PISA for identifying policy lessons. For example, in *World Class*, Schleicher (2018, 61) admitted that the results of PISA 'offer a snapshot of education at a certain moment in time', but 'cannot show how the school systems got to that point, or the institutions and organisations that might have helped or hindered progress'. He noted that 'the data do not really say anything about cause and effect', and 'knowing what successful systems are doing does not tell us how to improve less-successful systems' (ibid.).

Accordingly, to address these limitations, the OECD launched the Education 2030 Initiative (2018a) to help bring fresh ideas to PISA. When asked about how to maintain consistency in measuring improvements, my interview participant, however, stated that PISA prioritises relevance more than consistency:

> I know this is always the debate. Some people say, 'if you want to measure change, you cannot change the measure'. But we for PISA have always taken the attitude that it is more important to be relevant than to be consistent. … PISA had to evolve. [Otherwise] PISA would not be relevant; it would be consistent but not relevant. So, for PISA we always *prioritise to be relevant* to measure the skills that are needed today even though it means sometimes they are difficult to compare. (OECD Interview#2, 2018, my emphasis)

As these statements illustrate, the OECD has decided that the assessment regime should continuously evolve rather than function as a fixed point focusing on 'what was

considered important some point in the past' (Schleicher 2018, 277). In this regard, the 2030 Learning Compass (OECD 2019b) would be the rudder, allowing PISA to lead education reform by measuring students against what they *will* need to thrive in the future.

In the current world heavily affected by the global COVID-19 pandemic, it seems that the OECD's 'relevance first' strategy works well, as it does allow the Organisation to dismiss critiques as time-lapsed and, more crucially, to add fresh layers of data and ideas to rejuvenate the test. For example, the OECD has recently introduced several COVID-related questionnaires to assess government and school responses, teacher preparedness, and student well-being. In particular, the Global Crises Questionnaire Module aims to capture students' learning experiences during the lockdown, and the items will be incorporated into future PISA cycles. In the meantime, the OECD has released three additional volumes of PISA 2018 results, focusing on themes of 15-year-old students' understanding of money (Vol. IV: *Are Students Smart about Money?*), school policies and practices (Vol. V: *Effective Policies, Successful Schools*), and global competency (Vol. VI: *Are Students Ready to Thrive in an Interconnected World?*).

Accompanied by test specialists (or analysts), Schleicher has widely disseminated these new tools and volumes through live presentations and question-and-answer sessions. I note that Schleicher, in one way or another, has attempted to relate the emerging themes and issues with existing PISA data, especially when addressing questions about the relevance of the data collected prior to the pandemic. Thus, it would appear that the OECD is not only attempting to renew the test but also wants to maintain the connection with existing data sets for the purpose of interpretation and providing effective policy solutions.

Although the OECD highlights the importance of including out-of-school youth in PISA, it has ignored this group from these recent initiatives. This group of young people – who are the most affected by the pandemic and will be left further behind – were not taken into account when, for example, the OECD introduced the PISA Global Crisis Module and the Global Competency component (both school-based instruments). When I asked about why they did not consider out-of-school settings, Schleicher had no excuse:

> It is a good question. It is a fair question. That is a *real limitation* of PISA as it stands today that it does *not* include students who are not enrolled in school. Basically, a school-based instrument is a school-based test. … So, I think, that is something PISA *should* be doing but it is just operationally very, very difficult. We have actually made a start with this, last year, with a few countries developing an out-of-school test for PISA, so that we actually collected information on 15-year-olds who were not in school, and the results are going to be published also before the end

of this year. But that's just, sort of more of an experiment to fill that gap. But I think it is a very legitimate question, and a limitation of PISA as it is today *just* looking at those who are enrolled in school.

<div align="right">(Schleicher's answer to my question during the Webinar promoting the PISA 2018 results on global competency, 20 October 2020, original emphasis)</div>

Schleicher's acknowledgement of the 'real limitation' of PISA, which has not been resolved through the out-of-school component attempted in PISA-D, undermines the declared experimental success of this project (see also below). It further challenges the positioning of PISA as a barometer of inclusion that it has recently gained through supporting learning for all. The newly published report on the results of PISA-D out-of-school assessment stated that,

> […] in many low- and middle-income countries, including some of those that have participated in PISA, relatively large proportions of 15-year-olds are not enrolled in school or are not enrolled in PISA's target grades (grade 7 and above). Today, […] there are 61 million children of lower secondary school age out of school around the world. Through PISA-D, the OECD set out to ensure that this population is no longer beyond the reach of programmes that focus on evaluating the readiness of young people for their full participation in society. (OECD 2020, 2)

However, as examined earlier in Chapter 6, it is at best problematic that the OECD defined children who are enrolled in school but have repeated a grade and are below PISA's target grades as being 'out-of-school'. While contrary to the UIS and UNICEF definitions of that group, it does allow the OECD to incorporate a sub-population – which accounts for 22 percent on average across the five participating countries – and, therefore, provide access to readily available data derived from those in school. In effect, the notion of out-of-school youth has been defined to match what is easily measurable; an approach also employed in the OECD's measurement of global competency.

This situation echoes the literature which considers testing and inclusion as two competing global education policy agendas (Rizvi and Lingard 2006, 2009; Hardy and Woodcock 2015; Hamre, Morin and Ydesen 2018). A recent paper by Ydesen et al. (2020) highlights the 'capacity' of testing to exclude those who do not obtain the 'right' knowledge in the 'right' way. In the case of the OECD, it appears that PISA furthers the exclusion of out-of-school youth as it continues to produce new and relevant testing tools without accounting for this group, and it is unwilling to tackle the operational challenges that PISA-D encountered during the implementation phases.

I suggest that the OECD acts as an empire, one whose boundaries are not temporary. In addition to expanding its frontiers of operation, the empire prioritises relevance over

consistency and keeps reinventing the test rather than spending time and energy solving the really important problems. By highlighting the OECD's failure and unwillingness to address the 'real limitation' of PISA, this thesis challenges the authority of its expertise and the legitimacy of its moral claims.

Limitations and areas for future research

This thesis analyses the OECD's extension of PISA into low- and middle-income countries – PISA for development – by tracing the origins, highlighting the legitimation strategies, and examining the outcomes. The interpretive process has been guided by a critical hermeneutic approach, which entails a strong involvement of values and an iterative–recursive mode of abductive reasoning. Reflexivity is another hallmark of this approach; it concerns the researcher's self-conscious reflections on her positionality, choice of data generation methods, and ways of interpretation. I have discussed these dimensions quite extensively in the methodology chapter and applied them in the analysis chapters. In this chapter, these aspects are further engaged as I summarise the previous chapters and my contributions to knowledge. Reflexivity also means to acknowledge the limitations of the study and to point out areas for future research. While this thesis has achieved its aims, it has some clear limitations. I reflect on them below.

The first limitation relates to the *scope* of this research project. The initial research design intended to cover two additional research questions: one focused on the GEP field and the other concerned with the local. Specifically, I wanted to understand the dynamic relations between the OECD, the World Bank, UNESCO and UNICEF over the leadership in measuring educational quality (i.e. SDG 4); I also wanted to bring in the voices of PISA-D participating countries with respect to their rationales for participation and the specific issues and challenges that they had confronted. I did not wish to support a binary view of the global and the local, as indicated by the two major theoretical positions – world culture vs. local agency – in the field of CE. Rather, in line with Auld, Morris and Rappleye (2019), I wanted to combine 'granular' empirical work in specific countries with attention to the changing rationales and agendas of IOs.

However, these questions turned out to be too broad and unfeasible for two reasons. On the one hand, in order to better understand the relations between these key agencies in measuring SDG 4, I would need to understand how they interacted in the past, how

their relations developed over time, and the OECD's role and position within all these. As noted in Chapter 4, although the OECD came to play a prominent role in the formation and monitoring of SDG 4, its footing in the field of EID can be loosely traced back to the early 1960s (e.g. the DAC; see also Mundy 1998). This meant that a deeper understanding can only be achieved by reviewing the major educational activities of each of these agencies from their inception. In fact, in the second year, I had tentatively traced the trajectories of these agencies. As shown in Table 7.1, I focused on themes such as mandate, approach to education, key transitions, and changing roles, which led up to the post-2015 era. But I soon realised that this type of analysis itself could be a separate research project, for example, a historical study based on a competent conceptual framework.

Considering the length of time designated to this aspect and the lack of a suitable framework, I decided to narrow down and mainly focus on the OECD. Narrowing the scope also enabled me to set a solid basis for interpreting the legitimation strategies and PISA-D outcomes. However, I have tried to allude to these dynamics in Chapters 1 and 5. In particular, in Chapter 5, I showed how the OECD succeeded in gaining moral authority by firstly invoking and affiliating with UNESCO and UNICEF and then capitalising on the established 'collaborative' relations to promote PISA-D. The tension was obvious, especially when the UIS GEM Report team issued a clarification and accused the OECD of presenting a one-sided story. Having developed a deeper understanding of the OECD's changing agendas and approaches to education and development (Chapter 4), along with organisational legitimacy as a useful heuristic (Chapter 5), I will continue with this aspect of inquiry in the near future. I believe that this would contribute to the literature on the changing relationship between key players in the GEP field.

Table 7.1 Overview of histories of the key traditional agencies in the field of EID and the OECD's involvement.

Key Agencies	THE WORLD BANK		UNESCO	unicef
Establishment and mandate	1944 Poverty reduction		1946 Illiteracy reduction (peace)	1946 Emergency relief (children)
Approach	Lending temporary loans (financial)		Holding conference (intellectual)	Supporting health & nutrition (material)
Cold War period Education agenda	Since the 1960s Economic approach: manpower forecasting, technical and vocational training		Key reference for other IOs Rights-based approach: freedom, humanity and rights	Since the 1960s Basic needs approach: child and maternal care
Changing role and approach	By the 1980s Took over UNESCO Structural adjustment policies		The 1970s and 1980s **In crisis!**	By the 1980s Broke up with UNESCO Address education setbacks
The transitional 1990s **EFA** movement: Why involved?	Alleviate effects of structural adjustment		Regain legitimacy	Raise profile in education Took human rights approach
Level of focus	Primary education		A wider consideration	Girls' education
Since the 2000s MDGs (on access, primary)	Agenda setter (dominant)		Putative leader (GMR)	Agenda setter
OECD came to become prominent				
Post-2015 context SDGs (on quality, secondary)	Prominent PISA & PISA-D Learning Framework 2030	Dominant WBR 2018 Learning to realise education's promise	Custodian agency[11] UIS as the official source of cross-national comparative data on education (GAML) and GEM Report	Co-/custodian agency Regional assessments and SEA-PLM

---

[11] The custodian agencies work with national statistical offices to develop methodologies for indicators to measure progress on the SDGs. The agencies also work with countries to compile data for SDG indicators, which they submit to the UN Statistics Global SDG database.

On the other hand, in the case of local specificities, I have found a rich body of literature which investigates the multiple reasons that drive countries' participation, and which raises concerns and criticisms of the extension of ILSAs to low-and middle-income countries. The OECD was also aware of the challenges and commissioned (former) OECD and World Bank employees (Bloem 2013; Lockheed, Prokic-Breuer and Shadrova 2015) to conduct systematic reviews of these countries' experience in engaging in PISA. Therefore, in a way, PISA-D was the solution to these problems. I have thus decided to *examine* the outcomes, focusing on how PISA-D has (or has not) addressed the challenges that low- and middle-income countries face, how the results are generated and compared, and what insights they provide for the participating countries. This decision was taken not to neglect the 'granular' empirical work, but rather to examine meticulously the outcomes which the OECD declared as successful and are essential for establishing its role in monitoring SDG 4. In Chapters 5 and 6, I have tried to incorporate some empirical work in specific cases, such as Cambodia, Brazil and Vietnam, to challenge the OECD's dominant portrayal. I did this by intertextually reading different source materials where the realities unfolded.

The second limitation of this research relates to *theorising*. I have illuminated the OECD's emerging governance approach which configures its multidisciplinary expertise and a governance toolkit and applies them globally. In education, this manifests in a larger PISA programme to incorporate developing countries and to monitor SDG 4. In addition to enlarging the scale and scope, it also delves into curriculum and pedagogy, by defining skills needed in the future and by aligning PISA test items with teaching and learning content. While I interpret this emerging approach as a form of 'configuration' control, there is some degree of ambiguity about this concept and its changing nature. I suggest that this is a highly advanced mode of governing: compared to existing technologies, it has both a horizontal dimension of coverage, diversity, inclusiveness and flexibility, and a vertical dimension of depth, integration, alignment and individuality.

Some scholars have attempted to capture and conceptualise this current state, for example, as the new 'paradigm' for development (Auld and Morris 2019; Auld, Morris and Rappleye 2019), a spreading 'global web of measurement' (Moss and Urban 2019), a kind of 'ontological flattening' (Addey and Gorur 2020), topological 'respatialisations' (Lewis 2018, 2020), and a 'humanitarian turn' (Li and Auld 2020). This emerging approach has certainly taken on new features, but the underlying logic – a never-ending

future orientation and constant adaptation – seems to remain consistent. Indeed, as the OECD (2012d, 3) itself acknowledges, its strategy is 'evolutionary in nature, cognisant of the need for the Organisation's work on development to continuously evolve to be more responsive to the global realities and development needs, and for the internal mechanisms and external partnerships to deliver accordingly'.

From these factors, it would appear that the OECD's evolving strategy features uncertainty at its core but that is necessary for ensuring organisational survival and continual expansion. As I have argued, the OECD adapts as it expands, but this expansion has limits. As such, it is constantly gaining, maintaining and repairing its legitimacy. Similarly, the OECD's configuration of expertise and toolkit is unresolved, with some elements already coming into operation and others under development, such as surveys of student learning experiences and well-being, and the integration of new technologies to help upgrade its analytical capacity. The OECD's CERI project on 'Smart Data and Digital Technology in Education: Learning Analytics, AI and beyond'[12], which started before the pandemic, explores how digitalisation transforms the education sector and the learning processes through the use of data.

The current global learning crisis is accelerating these new frontiers of development. However, if the OECD wants to maintain the role of the 'central cog' within this nascent digital governance (Williamson 2017), it will have to ensure that the education GPS is in its own hand. In the near future, I will examine how the OECD interacts with education technology companies (as the 'new powers'; Heimans and Timms 2018) and how they affect the future education agenda as debates around replacing the SDGs will gather pace over the next five years.

As noted in the research design, there is no 'conclusion' in the interpretive research cycle; only momentary stopping points to collect one's thoughts and to move forward in a continuing process towards a deeper and richer understanding. I have, at certain points, stopped to collect my thoughts and share and work with colleagues to publish what I understood at that point in time (Li and Auld 2020; Auld, Li and Morris 2020) and then continued along this path. This concluding chapter is part of the journey leading towards a richer understanding.

---

[12] Accessed on 11 November 2020, https://www.oecd.org/education/ceri/smart-data-digital-technology-education-learning-analytics-ai.htm

References

Abbott, Kenneth W., and Duncan Snidal. 1998. 'Why States Act through Formal International Organisations'. *The Journal of Conflict Resolution* 42 (1): 3–32.

Addey, Camilla. 2016a. 'Camilla Addey Interviews Andreas Schleicher on PISA and the OECD'. Laboratory of International Assessment Studies. 2016. http://international-assessments.org/camilla-addey-interviews-andreas-schleicher-on-pisa-and-pisa-for-development/.

———. 2016b. 'PISA for Development and the Sacrifice of Policy-Relevant Data'. *Educação & Sociedade* 37 (136): 685–706. https://doi.org/10.1590/es0101-73302016166001.

———. 2017. 'Golden Relics & Historical Standards: How the OECD Is Expanding Global Education Governance through PISA for Development'. *Critical Studies in Education* 0 (0): 1–15. https://doi.org/10.1080/17508487.2017.1352006.

———. 2019. 'The Appeal of PISA for Development in Ecuador and Paraguay: Theorising and Applying the Global Ritual of Belonging'. *Compare: A Journal of Comparative and International Education* 0 (0): 1–16. https://doi.org/10.1080/03057925.2019.1623653.

Addey, Camilla, and Radhika Gorur. 2020. 'Translating PISA, Translating the World'. *Comparative Education* 0 (0): 1–18. https://doi.org/10.1080/03050068.2020.1771873.

Addey, Camilla, and Sam Sellar. 2018. 'Why Do Countries Participate in PISA? Understanding the Role of International Large-Scale Assessments in Global Education Policy'. In *Global Education Policy and International Development: New Agendas, Issues and Policies*, by Antoni Verger, Mario Novelli, and Hülya K. Altinyelken, 2nd ed. London: Bloomsbury Academic.

———. 2019. 'Is It Worth It? Rationales for (Non)Participation in International Large-Scale Learning Assessments'. 24. Education Research and Foresight Working Papers Series. Paris: UNESCO.

Agnese, Vasco d'. 2015. 'PISA's Colonialism: Success, Money, and the Eclipse of Education'. *Power and Education* 7 (1): 56–72. https://doi.org/10.1177/1757743814567387.

Aikenhead, Glen S. 1997. 'A Framework for Reflecting on Assessment and Evaluation'. In Globalisation of Science Education - Seoul International Conference. 195–99. Seoul: Korean Education Development Institute.

Allen, Mike. 2017. 'Interpretive Research'. In *The SAGE Encyclopedia of Communication Research Methods*. Thousand Oaks: SAGE Publications. https://doi.org/10.4135/9781483381411.n273.

Arkinson, Paul A., and Amanda Coffey. 2004. 'Analysing Documentary Realities'. In *Qualitative Research: Theory, Method and Practice*, edited by David Silverman, 2nd ed., 56–75. London: SAGE Publications.

———. 2011. 'Analysing Documentary Realities'. In *Qualitative Research: Issues of Theory, Method and Practice*, edited by David Silverman, 3rd ed., 77–92. London: SAGE Publications.

Arnett, Jeffrey J. 2008. 'The Neglected 95%: Why American Psychology Needs to Become Less American'. *The American Psychologist* 63 (7): 602–14. https://doi.org/10.1037/0003-066X.63.7.602.

Auld, Euan, Xiaomin Li, and Paul Morris. 2020. 'Piloting PISA for Development to Success: An Analysis of Its Findings, Framework and Recommendations'. *Compare: A Journal of Comparative and International Education* 0 (0): 1–25. https://doi.org/10.1080/03057925.2020.1852914.

Auld, Euan, and Paul Morris. 2014. 'Comparative Education, the "New Paradigm" and Policy Borrowing: Constructing Knowledge for Educational Reform'. *Comparative Education* 50 (2): 129–55. https://doi.org/10.1080/03050068.2013.826497.

———. 2019. 'Science by Streetlight and the OECD's Measure of Global Competence: A New Yardstick for Internationalisation?' *Policy Futures in Education* 17 (6): 677–98. https://doi.org/10.1177/1478210318819246.

Auld, Euan, Jeremy Rappleye, and Paul Morris. 2019. 'PISA for Development: How the OECD and World Bank Shaped Education Governance Post-2015'. *Comparative Education* 55 (2): 197–219. https://doi.org/10.1080/03050068.2018.1538635.

Avvisati, Francesco. 2019. 'Enhancing Cognitive Instruments for PISA-D'. Presented at the PISA for Development International Seminar, London, September 25 2019. http://www.oecd.org/pisa/pisa-for-development/4._PISA-D_Seminar_2019_Avvisati.pdf.

Baird, Jo-Anne, David Andrich, Therese N. Hopfenbeck, and Gordon Stobart. 2017. 'Assessment and Learning: Fields Apart?' *Assessment in Education: Principles, Policy & Practice* 24 (3): 317–50. https://doi.org/10.1080/0969594x.2017.1319337.

Ball, Stephen J. 2003. 'The Teacher's Soul and the Terrors of Performativity'. *Journal of Education Policy* 18 (2): 215–28. https://doi.org/10.1080/0268093022000043065.

———. 2012. *Global Education Inc.: New Policy Networks and the Neo-Liberal Imaginary*. New York: Routledge. https://www.dawsonera.com/abstract/9780203803301.

Bandura, Albert. 1969. *Principles of Behavior Modification*. New York: Holt, Rinehart and Winston.

Barnett, Michael, and Martha Finnemore. 2004. *Rules for the World: International Organisations in Global Politics*. New York: Cornell University Press. https://www.jstor.org/stable/10.7591/j.ctt7z7mx.

Barnett, Michael N., and Martha Finnemore. 1999. 'The Politics, Power, and Pathologies of International Organisations'. *International Organisation* 53 (4): 699–732. https://doi.org/10.1162/002081899551048.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 1st ed. National Bureau of Economic Research. General Series; No. 80. New York: National Bureau of Economic Research.

Becker, Howard S. 1998. *Tricks of the Trade: How to Think about Your Research While You're Doing It*. Chicago Guides to Writing, Editing, and Publishing. Chicago: University of Chicago Press.

———. 2007. *Writing for Social Scientists: How to Start and Finish Your Thesis, Book, or Article*. 2nd ed. Chicago Guides to Writing, Editing, and Publishing. Chicago: University of Chicago Press.

Benavot, Aaron, and Heniz-Dieter Meyer. 2013. *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Oxford: Symposium Books.

Bengtsson, Jarl. 2008. 'OECD's Centre for Educational Research and Innovation—1968 to 2008'. In *Learning in the 21st Century: Research, Innovation and Policy*, 4. Paris: OECD Publishing.

Bentz, Valerie M., and Jeremy J. Shapiro. 1998. *Mindful Inquiry into Social Research*. Thousand Oaks: SAGE Publications.

Berry, Jeffrey M. 2002. 'Validity and Reliability Issues in Elite Interviewing'. *Political Science & Politics* 35 (4): 679–82. https://doi.org/10.1017/S1049096502001166.

Biesta, Gert. 2008. 'Good Education in an Age of Measurement: On the Need to Reconnect with the Question of Purpose in Education'. *Educational Assessment, Evaluation and Accountability (Formerly: Journal of Personnel Evaluation in Education)* 21 (1): 33–46. https://doi.org/10.1007/s11092-008-9064-9.

Billig, Michael. 1995. *Banal Nationalism*. Los Angeles: SAGE Publications.

Bitektine, Alex. 2011. 'Toward a Theory of Social Judgments of Organisations: The Case of Legitimacy, Reputation, and Status'. *The Academy of Management Review* 36 (1): 151–79.

Bloem, Simone. 2013. 'PISA in Low- and Middle-Income Countries'. OECD Education Working Papers. Paris: OECD Publishing. http://www.oecd-ilibrary.org/content/workingpaper/5k41tm2gx2vd-en.

———. 2015. 'The OECD Directorate for Education as an Independent Knowledge Producer through PISA'. In *Governing Educational Spaces: Knowledge, Teaching, and Learning in Transition*, edited by Hans-Georg Kotthoff and Eleftherios Klerides, 169–88. Comparative Education Society in Europe Series. Rotterdam: Sense Publishers.

Boel, Bent. 2003. *The European Productivity Agency and Transatlantic Relations, 1953-1961*. Studies in 20th and 21st Century European History 4. Copenhagen: Museum Tusculanum Press.

Bogner, Alexander, Beate Littig, and Wolfgang Menz, eds. 2009. *Interviewing Experts*. Research Methods Series. Basingstoke: Palgrave Macmillan. http://www.vlebooks.com/vleweb/product/openreader?id=UCL&isbn=97802302 44276.

Bourdieu, Pierre. 1992. *An Invitation to Reflexive Sociology*. Cambridge: Polity Press.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. 'The Determinants of Earnings: A Behavioral Approach'. *Journal of Economic Literature* 39 (4): 1137–76. https://doi.org/10.1257/jel.39.4.1137.

Braun, Henry, and Anil Kanjee. 2006. 'Using Assessment to Improve Education in Developing Nations'. In *Improving Education through Assessment, Innovation,*

and Evaluation, edited by Henry Braun, Anil Kanjee, Eric Bettinger, and Michael Kremer, 1–46. Cambridge: American Academy of Arts and Sciences.

Brock-Utne, Birgit. 2016. 'The Ubuntu Paradigm in Curriculum Work, Language of Instruction and Assessment'. *International Review of Education* 62 (1): 29–44. https://doi.org/10.1007/s11159-016-9540-2.

Bryman, Alan. 2012. *Social Research Methods*. 4th ed. Oxford: Oxford University Press.

Buchanan, Allen, and Robert O. Keohane. 2006. 'The Legitimacy of Global Governance Institutions'. *Ethics and International Affairs* 20 (4): 405–37. https://doi.org/10.1111/j.1747-7093.2006.00043.x.

Bürgi, Regula. 2019. 'Learning Productivity: The European Productivity Agency—An Educational Enterprise'. In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 17–38. Global Histories of Education. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-33799-5.

Buzzanell, Patrice M. 2018. 'Legitimising and Transforming the Closet/Closeting'. *Management Communication Quarterly* 32 (2): 297–300. https://doi.org/10.1177/0893318917742518.

Carney, Stephen, Jeremy Rappleye, and Iveta Silova. 2012. 'Between Faith and Science: World Culture Theory and Comparative Education'. *Comparative Education Review* 56 (3): 366–93. https://doi.org/10.1086/665708.

Carnoy, Martin, and Richard Rothstein. 2015. 'What International Test Scores Tell Us'. *Society (New Brunswick)* 52 (2): 122–28. https://doi.org/10.1007/s12115-015-9869-3.

Carr, Edward H. 1964. *The Twenty Years' Crisis, 1919-1939: An Introduction to the Study of International Relations*. New York: Harper and Row.

Carr-Hill, Roy. 2013. 'Missing Millions and Measuring Development Progress'. *World Development* 46 (June): 30–44. https://doi.org/10.1016/j.worlddev.2012.12.017.

———. 2015. 'PISA for Development Technical Strand C: Incorporating out-of-School 15-Year-Olds in the Assessment'. OECD Education Working Papers 120. Paris: OECD Publishing. http://dx.doi.org/10.1787/5js0bsln9mg2-en.

Centeno, Vera G. 2019. 'The Birth of the OECD's Education Policy Area'. In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 63–82. Global Histories of Education. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-33799-5.

Coetzee, John M. 2004. *Waiting for the Barbarians*. London: Vintage.

Coffey, Amanda. 2013. 'Analysing Documents'. In *The SAGE Handbook of Qualitative Data Analysis*, edited by Uwe Flick, 367–79. Thousand Oaks: SAGE Publications. http://orca.cf.ac.uk/53901/.

Corbin, Juliet M., and Anselm L. Strauss. 2008. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 3rd ed. London: SAGE Publications.

Cowen, Robert. 2006. 'Acting Comparatively upon the Educational World: Puzzles and Possibilities'. *Oxford Review of Education* 32 (5): 561–73. https://doi.org/10.1080/03054980600976155.

———. 2007. 'Comparing and Transfering: Visions, Politics and Universities'. In *Higher Education and National Development: Universities and Societies in Transition*, edited by David Bridges, Palmira Juceviciene, Roberta Jucevicius, Terence H. Mclaughlin, and Jolanta Stankeviciute, 13–29. Oxon: Routledge. https://books.google.com/books/about/Higher_Education_and_National_Develo pmen.html?id=g0CAAgAAQBAJ.

———. 2014. 'Ways of Knowing, Outcomes and "Comparative Education": Be Careful What You Pray For'. *Comparative Education* 50 (3): 282–301. https://doi.org/10.1080/03050068.2014.921370.

———. 2018. 'Comparative Education and Empires'. *Revista Española de Educación Comparada*, no. 31: 14–34. https://doi.org/10.5944/reec.31.2018.21828.

Crossley, Michael. 2014. 'Global League Tables, Big Data and the International Transfer of Educational Research Modalities'. *Comparative Education* 50 (1): 15–26. https://doi.org/10.1080/03050068.2013.871438.

Crouch, Luis, and Silvia Montoya. 2018. 'SDG 4 Data: Investing Millions Today Will Save Billions in the Future'. UNESCO Institute for Statistics. 2018. http://uis.unesco.org/en/blog/sdg-4-data-investing-millions-today-will-save-billions-future.

Dahlberg, Gunilla, Peter Moss, and Alan Pence. 2013. *Beyond Quality in Early Childhood Education and Care: Languages of Evaluation*. Oxon: Routledge. https://doi.org/10.4324/9780203371114.

Davidson, Michael, Michael Ward, and Alejandro G. Palma. 2014. 'PISA for Development and the Post-2015 Agenda'. *NORRAG* (blog). 2014. https://www.norrag.org/pisa-for-development-and-the-post-2015-agenda/.

Deetz, Stanley. 1996. 'Describing Differences in Approaches to Organisation Science: Rethinking Burrell and Morgan and Their Legacy'. *Organisation Science* 7 (2): 191–207.

Diener, Edward, and Rick Crandall. 1978. *Ethics in Social and Behavioural Research*. Chicago: University of Chicago Press. https://books.google.com/books/about/Ethics_in_Social_and_Behavioral_Resear ch.html?id=dMZqvgAACAAJ.

Dilthey, Wilhelm. 1976. *Selected Writings*. Edited and translated by H. P. Rickman. Cambridge: Cambridge University Press.

Dowling, John, and Jeffrey Pfeffer. 1975. 'Organisational Legitimacy: Social Values and Organisational Behaviour'. *Pacific Sociological Review* 18 (1): 122–36. https://doi.org/10.2307/1388226.

Eccleston, Richard. 2011. 'The OECD and Global Economic Governance'. *Australian Journal of International Affairs* 65 (2): 243–55. https://doi.org/10.1080/10357718.2011.550106.

Edwards, Brent D., Taeko Okitsu, Romina da Costa, and Yuto Kitamura. 2018. 'Organizational Legitimacy in the Global Education Policy Field: Learning from

UNESCO and the Global Monitoring Report'. *Comparative Education Review* 62 (1): 31–63. https://doi.org/10.1086/695440.

Edwards, Sophie. 2019. 'India's Re-Entry to PISA Triggers Mixed Response'. *Devex*, 20 February 2019, sec. Education. https://www.devex.com/news/sponsored/india-s-re-entry-to-pisa-triggers-mixed-response-94286.

Eide, Kjell. 1990. *30 Years of Educational Collaboration in the OECD*. Paris: OECD Publishing.

Eisenhower, Dwight D. 1958. 'Recommendations Relative to Our Educational System'. *Science Education* 42 (2): 103–6. https://doi.org/10.1002/sce.3730420203.

Elfert, Maren. 2018. *UNESCO's Utopia of Lifelong Learning: An Intellectual History*. Routledge Research in Lifelong Learning and Adult Education Series. New York: Routledge. https://doi.org/10.4324/9781315278131.

———. 2019. 'The OECD, American Power and the Rise of the "Economics of Education" in the 1960s'. In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 39–62. Global Histories of Education. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-33799-5.

Entman, Robert M. 2007. 'Framing Bias: Media in the Distribution of Power'. *Journal of Communication* 57 (1): 163–73. https://doi.org/10.1111/j.1460-2466.2006.00336.x.

Ertl, Hubert. 2006. 'Educational Standards and the Changing Discourse on Education: The Reception and Consequences of the PISA Study in Germany'. *Oxford Review of Education* 32 (5): 619–34.

Espeland, Wendy N., and Mitchell L. Stevens. 1998. 'Commensuration as a Social Process'. *Annual Review of Sociology* 24 (1): 313–43. https://doi.org/10.1146/annurev.soc.24.1.313.

———. 2008. 'A Sociology of Quantification'. *European Journal of Sociology* 49 (3): 401–36. https://doi.org/10.1017/S0003975609000150.

Ettelt, Stefanie, Nicholas Mays, and Pauline Allen. 2015. 'The Multiple Purposes of Policy Piloting and Their Consequences: Three Examples from National Health and Social Care Policy in England'. *Journal of Social Policy* 44 (2): 319–37. https://doi.org/10.1017/S0047279414000865.

Examinations Council of Zambia. 2018. 'Education in Zambia: Findings from Zambia's Experience in PISA for Development'. PISA for Development. Lusaka: Ministry of General Education. https://www.oecd.org/pisa/pisa-for-development/Zambia_PISA_D_national_report.pdf.

Fanon, Frantz. 1967. *The Wretched of the Earth*. Translated by Jean-Paul Sartre. Harmondsworth: Penguin.

Feniger, Yariv, and Michael Atia. 2018. 'Rethinking Cause and Effect: Analysing Economic Growth and PISA Scores over a Period of 15 Years'. In *World Yearbook of Education 2019: Comparative Methodology in the Era of Big Data and Global Networks*, edited by Radhika Gorur, Sam Sellar, and Gita Steiner-Khamsi, 104–19. London: Routledge. https://doi.org/10.4324/9781315147338.

Finnemore, Martha. 1993. 'International Organizations as Teachers of Norms: The United Nations Educational, Scientific, and Cultural Organisation and Science Policy'. *International Organisation* 47 (4): 565–97. https://doi.org/10.1017/S0020818300028101.

———. 1996a. 'Constructing Norms of Humanitarian Intervention'. In *The Culture of National Security: Norms and Identity in World Politics*, edited by Peter J. Katzenstein, 153–85. New Directions in World Politics. New York: Columbia University Press.

———. 1996b. 'Norms, Culture, and World Politics: Insights from Sociology's Institutionalism'. *International Organization* 50 (2): 325–47. https://doi.org/10.1017/S0020818300028587.

Forster, Nick. 1994. 'The Analysis of Company Documentation'. In *Qualitative Methods in Organisational Research: A Practical Guide*, edited by Catherine Cassell and Gillian Symon, 1st ed., 147–66. London: SAGE Publications.

Foucault, Michel. 1972. *The Archaeology of Knowledge*. Translated by Smith Sheridan. World of Man. London: Tavistock Publications.

———. 1979. *Discipline and Punish: The Birth of the Prison*. Translated by Alan Sheridan. New York: Vintage Books. http://www.aspresolver.com/aspresolver.asp?SOTH;S10021788.

Frederiksen, Norman. 1984. 'The Real Test Bias: Influences of Testing on Teaching and Learning'. *The American Psychologist* 39 (3): 193–202. https://doi.org/10.1037/0003-066X.39.3.193.

Gadamer, Hans-Georg. 1976. *Philosophical Hermeneutics*. Edited and translated by David R. Linge. Berkeley: University of California Press.

Geertz, Clifford. 1973. *The Interpretation of Cultures: Selected Essays*. New York: Basic Books. http://hdl.handle.net/2027/heb.01005.

Gemelli, Giuliana. 1996. 'American Influence on European Management Education: The Role of the Ford Foundation'. In *Management, Education and Competitiveness: Europe, Japan and the United States*, edited by Rolv P. Amdam, 1st ed., 38–68. London: Routledge.

Gilpin, Robert. 1981. *War and Change in World Politics*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511664267.

Gorur, Radhika. 2019. 'Old Power, New Power and Ontological Flattening: The Global "data Revolution" in Education'. In *New Practices of Comparison, Quantification and Expertise in Education: Conducting Empirically Based Research*, edited by Christina Mølstad and Daniel Pettersson, 66–82. Oxon: Routledge.

Gorur, Radhika, Estrid Sørensen, and Bryan Maddox. 2019. 'Standardising the Context and Contextualising the Standard: Translating PISA into PISA-D'. In *Science, Numbers and Politics*, edited by Markus J. Prutsch, 301–29. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-11208-0_14.

Grek, Sotiria. 2009. 'Governing by Numbers: The PISA "Effect" in Europe'. *Journal of Education Policy* 24 (1): 23–37. https://doi.org/10.1080/02680930802412669.

———. 2013. 'Expert Moves: International Comparative Testing and the Rise of Expertocracy'. *Journal of Education Policy* 28 (5): 695–709. https://doi.org/10.1080/02680939.2012.758825.

Grey, Sue, and Paul Morris. 2018. 'PISA: Multiple "Truths" and Mediatised Global Governance'. *Comparative Education* 54 (2): 109–31. https://doi.org/10.1080/03050068.2018.1425243.

Grotlüschen, Anke. 2018. 'Global Competence – Does the New OECD Competence Domain Ignore the Global South?' *Studies in the Education of Adults* 50 (2): 185–202. https://doi.org/10.1080/02660830.2018.1523100.

Gurría, Angel. 2017. 'Bridging the Data Divide for Development'. Speeches. OECD Secretary-General. 2017. https://www.oecd.org/about/secretary-general/bridging-the-data-divide-for-development.htm.

Habermas, Jürgen. 1974. *Theory and Practice*. Translated by John Viertel. London: Heinemann.

Hamilton, Mary. 2018. 'Contributing to the Common Good? Media Coverage of the International Largescale Assessment of Adult Skills (PIAAC) in Four National Contexts'. *Studies in the Education of Adults* 50 (2): 167–84. https://doi.org/10.1080/02660830.2018.1523101.

Hamre, Bjørn, Anne Morin, and Christian Ydesen. 2018. *Testing and Inclusive Schooling: International Challenges and Opportunities*. Routledge Research in International and Comparative Education. Oxon: Routledge.

Hans, Nicholas. 1958. *Comparative Education: A Study of Educational Factors and Traditions*. 1st ed. International Library of Sociology and Social Reconstruction. London: Routledge and Kegan Paul.

Hanushek, Eric A., and Ludger Woessmann. 2008. 'The Role of Cognitive Skills in Economic Development'. *Journal of Economic Literature* 46 (3): 607–68. https://doi.org/10.1257/jel.46.3.607.

Hardt, Michael, and Antonio Negri. 2001. *Empire*. Cambridge: Harvard University Press.

Hardy, Ian, and Stuart Woodcock. 2015. 'Inclusive Education Policies: Discourses of Difference, Diversity and Deficit'. *International Journal of Inclusive Education* 19 (2): 141–64. https://doi.org/10.1080/13603116.2014.908965.

Harvey, William S. 2011. 'Strategies for Conducting Elite Interviews'. *Qualitative Research* 11 (4): 431–41. https://doi.org/10.1177/1468794111404329.

Heimans, Jeremy, and Henry Timms. 2018. *New Power: How Power Works in Our Hyperconnected World – and How to Make It Work for You*. New York: Doubleday Books.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. 'The Weirdest People in the World?' *Behavioral and Brain Sciences* 33 (2–3): 61–83. https://doi.org/10.1017/S0140525X0999152X.

Henry, Miriam, Bob Lingard, Fazal Rizvi, and Sandra Taylor. 2001. *The OECD, Globalisation and Education Policy*. Vol. 13. Emerald Group Publishing Limited. https://books.emeraldinsight.com/page/detail/the-oecd-globalisation-

and-education-policy-m-henryThe-OECD,-Globalisation-and-Education-Policy/?k=9780080434490.

Holden, Laura, and Jeff Biddle. 2017. 'The Introduction of Human Capital Theory into Education Policy in the United States'. *History of Political Economy* 49 (4): 537–74. https://doi.org/10.1215/00182702-4296305.

Holsti, Kal J. 1985. *The Dividing Discipline: Hegemony and Diversity in International Theory*. Boston: Allen & Unwin.

Holsti, Ole R. 2004. 'Theories of International Relations'. In *Explaining the History of American Foreign Relations*, edited by Michael J. Hogan and Thomas G. Paterson, 2nd ed., 51–90. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511806445.005.

Hongler, Patricia. 2017. 'The Construction of a Western Voice: OECD and the First UNCTAD of 1964'. In *The OECD and the International Political Economy since 1948*, edited by Matthieu Leimgruber and Matthias Schmelzer, 137–58. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-60243-1.

Instituto Nacional de Evaluación Educativa. 2018. 'Educación en Ecuador: Resultados de PISA para el Desarrollo'. Asunción: Instituto Nacional de Evaluación Educativa. http://evaluaciones.evaluacion.gob.ec/BI/pisa/.

Jones, Phillip W., and David Coleman. 2005. *The United Nations and Education: Multilateralism, Development and Globalisation*. London: RoutledgeFalmer.

Junemann, Carolina, and Stephen Ball. 2015. 'Pearson and PALF: The Mutating Giant'. Education International. https://download.ei-ie.org/Docs/WebDepot/ei_palf_publication_web_pages_01072015.pdf.

Kaess, Kathleen. 2018. 'The Cross-Atlantic Knowledge Divide, or PISA for Development: Should One Size Ever Fit All?' *Atlantic Studies* 15 (3): 349–64. https://doi.org/10.1080/14788810.2017.1370356.

Kallo, Johanna. 2009. *OECD Education Policy: A Comparative and Historical Study Focusing on the Thematic Reviews of Tertiary Education*. Jyvaskyla: Finnish Educational Research Association.

Kamens, David H. 2013. 'Globalisation and the Emergence of an Audit Culture: PISA and the Search for "best Practices" and Magic Bullets'. In *PISA, Power, and Policy: The Emergence of Global Educational Governance*, edited by Heinz-Dieter Meyer and Aaron Benavot. Oxford: Symposium Books.

———. 2015. 'A Maturing Global Testing Regime Meets the World Economy: Test Scores and Economic Growth, 1960–2012'. *Comparative Education Review* 59 (3): 420–46. https://doi.org/10.1086/681989.

Kamens, David H., and Connie L. McNeely. 2010. 'Globalisation and the Growth of International Educational Testing and National Assessment'. *Comparative Education Review* 54 (1): 5–25. https://doi.org/10.1086/648471.

Kandel, Isaac L. 1955. *The New Era in Education: A Comparative Study*. Boston: Houghton Mifflin.

Kanjee, Anil. 2018. 'Use of International Assessments to Support Learning at the Bottom of the Pyramid'. In *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries*, edited by Daniel Wagner,

Sharon Wolf, and Robert F. Boruch, 254. Paris: International Institute for Educational Planning.

Kaplan, Abraham. 1964. *The Conduct of Inquiry: Methodology for Behavioural Science*. Chandler Publications in Anthropology and Sociology. San Francisco: Chandler PubCo.

Kazamias, Andreas M. 2009. 'Forgotten Men, Forgotten Themes: The Historical-Philosophical-Cultural and Liberal Humanist Motif in Comparative Education'. In *International Handbook of Comparative Education*, edited by Robert Cowen and Andreas M. Kazamias, 37–58. Springer International Handbooks of Education. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-6403-6_4.

Kelly, Nancy. 2002. 'Using Documentary Analysis to Reveal Narratives of Mothering'. In *Narrative, Memory and Life Transitions*, edited by Christine Horrocks, Kate Milnes, Brian Roberts, and David Robinson, 189–97. Huddersfield: University of Huddersfield. http://www2.hud.ac.uk/hhs/nme/books/2002/Chapter_20_Nancy_Kelly.pdf.

Kennedy, Ann. 2019. 'PISA for Development: Enhancement of Cognitive Instruments'. Presented at the PISA for Development International Seminar, London, September 15. http://www.oecd.org/pisa/pisa-for-development/5._PISA-D_Seminar_2019_Kennedy.pdf.

Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton: Princeton University Press.

———. 1988. 'International Institutions: Two Approaches'. *International Studies Quarterly* 32 (4): 379–96. https://doi.org/10.2307/2600589.

———. 2009. 'Political Science as a Vocation'. *Political Science & Politics* 42 (2): 359–63. https://doi.org/10.1017/S1049096509090489.

Kijima, Rie. 2013. 'The Politics of Cross-National Assessments: Global Trends and National Interests'. ProQuest Dissertations Publishing. https://search.proquest.com/docview/1459423842?pq-origsite=summon.

King, Kenneth. 2017. 'Lost in Translation? The Challenge of Translating the Global Education Goal and Targets into Global Indicators'. *Compare* 47 (6): 801–17. https://doi.org/10.1080/03057925.2017.1339263.

King, Nigel. 1994. 'The Qualitative Research Interview'. In *Qualitative Methods in Organisational Research: A Practical Guide*, edited by Catherine Cassell and Gillian Symon, 14–36. London: SAGE Publications.

Klees, Steven J. 2016. 'Human Capital and Rates of Return: Brilliant Ideas or Ideological Dead Ends?' *Comparative Education Review* 60 (4): 644–72. https://doi.org/10.1086/688063.

Klees, Steven J., Mark Ginsburg, Hiba Anwar, Melanie B. Robbins, Heidi Bloom, Cameron Busacca, Anne Corwith, et al. 2020. 'The World Bank's SABER: A Critical Analysis'. *Comparative Education Review* 64 (1): 46–65. https://doi.org/10.1086/706757.

Komatsu, Hikaru, and Jeremy Rappleye. 2017. 'A New Global Policy Regime Founded on Invalid Statistics? Hanushek, Woessmann, PISA, and Economic Growth'.

Comparative Education 53 (2): 166–91.
https://doi.org/10.1080/03050068.2017.1300008.

Krasner, Stephen D. 1982. 'Structural Causes and Regime Consequences: Regimes as Intervening Variables'. *International Organisation* 36 (2): 185–205. https://doi.org/10.1017/S0020818300018920.

Kvale, Steinar. 1983. 'The Qualitative Research Interview: A Phenomenological and a Hermeneutical Mode of Understanding'. *Journal of Phenomenological Psychology* 14 (2): 171-.

Leimgruber, Matthieu, and Matthias Schmelzer, eds. 2017. *The OECD and the International Political Economy since 1948*. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-60243-1.

Lewis, Steven. 2018. 'PISA "yet to Come": Governing Schooling through Time, Difference and Potential'. *British Journal of Sociology of Education* 39 (5): 683–97. https://doi.org/10.1080/01425692.2017.1406338.

———. 2020. *PISA, Policy and the OECD: Respatialising Global Educational Governance through PISA for Schools*. Springer Singapore. https://doi.org/10.1007/978-981-15-8285-1.

Li, Xiaomin, and Euan Auld. 2020. 'A Historical Perspective on the OECD's "Humanitarian Turn": PISA for Development and the Learning Framework 2030'. *Comparative Education* 56 (4): 503–21. https://doi.org/10.1080/03050068.2020.1781397.

Lincoln, Yavonna, and Egon G. Guba. 1985. *Naturalistic Inquiry*. Newbury Park: SAGE Publications.

Lingard, Bob. 2017. 'International Organisations' Rationales for Large Scale Assessments'. *Compare: A Journal of Comparative and International Education* 47 (3): 434–52. https://doi.org/10.1080/03057925.2017.1301399.

Lingard, Bob, and Shaun Rawolle. 2011. 'New Scalar Politics: Implications for Education Policy'. *Comparative Education* 47 (4): 489–502. https://doi.org/10.1080/03050068.2011.555941.

Lingard, Bob, Shaun Rawolle, and Sandra Taylor. 2005. 'Globalizing Policy Sociology in Education: Working with Bourdieu'. *Journal of Education Policy* 20 (6): 759–77. https://doi.org/10.1080/02680930500238945.

Lockheed, Marlaine. 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): 512–18. https://doi.org/10.1080/03057925.2012.670480.

———. 2015. 'Why Do Countries Participate in International Large-Scale Assessments?: The Case of PISA'. World Bank, Washington, DC.

———. 2018. 'Lockheed Commentary on Willms' Educational Prosperity: An Assessment Strategy for Supporting Student Learning in Low-Income Countries'. In *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries*, edited by Daniel A. Wagner, Sharon Wolf, and Robert F. Boruch. UNESCO-IIEP.

Lockheed, Marlaine E., and Hans Wagemaker. 2013. 'International Large-Scale Assessments: Thermometers, Whips or Useful Policy Tools?' *Research in Comparative and International Education* 8 (3): 296–306. https://doi.org/10.2304/rcie.2013.8.3.296.

Lockheed, Marlaine, Tijana Prokic-Bruer, and Anna Shadrova. 2015. 'The Experience of Middle-Income Countries Participating in PISA 2000–2015'. Paris: OECD Publishing. http://www.oecd-ilibrary.org/content/book/9789264246195-en.

Maddox, Bryan. 2014. 'Globalising Assessment: An Ethnography of Literacy Assessment, Camels and Fast Food in the Mongolian Gobi'. *Comparative Education* 50 (4): 474–89. https://doi.org/10.1080/03050068.2013.871440.

———. 2015. 'The Neglected Situation: Assessment Performance and Interaction in Context'. *Assessment in Education: Principles, Policy & Practice* 22 (4): 427–43. https://doi.org/10.1080/0969594X.2015.1026246.

Martens, Kerstin. 2007. 'How to Become an Influential Actor - The "comparative Turn" in OECD Education Policy'. In *New Arenas of Education Governance: The Impact of International Organisations and Markets on Educational Policy Making*, edited by Kerstin Martens, Alessandra Rusconi, and Kathrin Leuze. Transformations of the State ^A198299. Basingstoke: Palgrave Macmillan.

Martens, Kerstin, and Anja P. Jakobi. 2010. *Mechanisms of OECD Governance: International Incentives for National Policy-Making?* Oxford: Oxford University Press. https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199591145.001.0001/acprof-9780199591145.

Mattheou, Dimitris. 2009. 'The Scientific Paradigm in Comparative Education'. In *International Handbook of Comparative Education*, edited by Robert Cowen and Andreas M. Kazamias, 59–71. Springer International Handbooks of Education. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-6403-6_5.

Maxwell, Joseph A. 2009. 'Designing a Qualitative Study'. In *The SAGE Handbook of Applied Social Research Methods*, edited by Leonard Bickman and Debra Rog, 2nd ed., 214–53. Thousand Oaks: SAGE Publications. https://doi.org/10.4135/9781483348858.

McCann, Eugene. 2011. 'Urban policy mobilities and global circuits of knowledge: Toward a research agenda'. *Annals of the Association of American Geographers* 101 (1): 107–30. https://doi.org/10.1080/00045608.2010.520219.

McClelland, David C. 1961. *The Achieving Society*. New York: Van Nostrand.

McCowan, Tristan. 2015. 'Theories of Development'. In *Education and International Development: An Introduction*, edited by Tristan McCowan and Elaine Unterhalter, 14–31. London: Bloomsbury.

McCowan, Tristan, and Elaine Unterhalter. 2015. *Education and International Development: An Introduction*. London: Bloomsbury.

McKinlay, Robert D., and Richard Little. 1986. 'Global Problems and World Order'. *Verfassung Und Recht in Übersee* 19 (4): 495–96. https://doi.org/10.5771/0506-7286-1986-4-495.

Mehta, Lyla. 2001. 'The World Bank and Its Emerging Knowledge Empire'. *Human Organisation* 60 (2): 189–96. https://doi.org/10.17730/humo.60.2.cl4dw5yv0gv7867c.

Menashy, Francine. 2017. 'Multi-Stakeholder Aid to Education: Power in the Context of Partnership'. *Globalisation, Societies and Education* 16 (1): 13–26. https://doi.org/10.1080/14767724.2017.1356702.

Meyer, Heinz-Dieter, and Katie Zahedi. 2014. 'Open Letter to Andreas Schleicher, OECD, Paris'. *Policy Futures in Education* 12 (7): 872–77. https://doi.org/10.2304/pfie.2014.12.7.872.

Mikecz, Robert. 2012. 'Interviewing Elites: Addressing Methodological Issues'. *Qualitative Inquiry* 18 (6): 482–93. https://doi.org/10.1177/1077800412442818.

Milner, Helen. 1992. 'International Theories of Cooperation among Nations: Strengths and Weaknesses'. *World Politics* 44 (3): 466–96. https://doi.org/10.2307/2010546.

Mincer, Jacob A. 1974. *Schooling, Experience, and Earnings*. Cambridge: National Bureau of Economic Research. https://www.nber.org/books-and-chapters/schooling-experience-and-earnings.

Ministerio de Educación. 2018. 'Guatemala en PISA-D: Programa Internacional de Evaluación de Estudiantes'. PISA for Development. Guatemala: Dirección General de Evaluación e Investigación Educativa. http://www.mineduc.gob.gt/digeduca/documents/pisa/InformePISADGuatemala.pdf.

Ministerio de Educación y Ciencias. 2018. 'Educación en Paraguay: Hallazgos de la PISA para el Desarrollo'. Asunción: Ministerio de Educación y Ciencias. https://www.oecd.org/pisa/pisa-for-development/PISAD-Paraguay-InSchool-National-Report.pdf.

Ministre de l'Éducation nationale. 2018. 'L'Éducation au Sénégal: Résultats de l'enquete PISA-D 2017 au Sénégal'. Dakar: Ministre de l'Éducation nationale. http://www.oecd.org/pisa/pisa-for-development/Senegal_PISA_D_national_report.pdf.

MoEYE. 2018. 'Education in Cambodia: Findings from Cambodia's Experience in PISA for Development'. Phnom Penh. https://www.oecd.org/pisa/pisa-for-development/PISA-D%20national%20report%20for%20Cambodia.pdf.

Mølstad, Christina, and Daniel Pettersson. 2019. *New Practices of Comparison, Quantification and Expertise in Education: Conducting Empirically Based Research*. Oxon: Routledge.

Montoya, Silvia. 2019. 'Global Review of Progress towards SDG4-Education2030 Targets and Commitments'. Presented at the PISA for Development International Seminar, London, September 25 2019. http://www.oecd.org/pisa/pisa-for-development/19._PISA-D_Seminar_2019_Montoya.pdf.

Morgan, Clara. 2017. 'Problematizing the Expansion of PISA's Explanatory Power'. In *The PISA Effect on Global Educational Governance*, edited by Louis Volante, 17–31. New York: Routledge. https://doi.org/10.4324/9781315440521.

Morgenthau, Hans J. 1967. *Politics among Nations: The Struggle for Power and Peace*. 4th ed. New York: Alfred A Knopf.

Morris, Paul. 2012. 'Pick "n" Mix, Select and Project; Policy Borrowing and the Quest for "World Class" Schooling: An Analysis of the 2010 Schools White Paper'. *Journal of Education Policy* 27 (1): 89–107. https://doi.org/10.1080/02680939.2011.596226.

———. 2015. 'Comparative Education, PISA, Politics and Educational Reform: A Cautionary Note'. *Compare: A Journal of Comparative and International Education* 45 (3): 470–74. https://doi.org/10.1080/03057925.2015.1027510.

———. 2016. *Education Policy, Cross-National Tests of Pupil Achievement, and the Pursuit of World-Class Schooling: A Critical Analysis*. Professorial Lecture Series (University College London. Institute of Education). London: UCL Institute of Education Press.

Moss, Peter. 2016. 'Why Can't We Get beyond Quality?' *Contemporary Issues in Early Childhood* 17 (1): 8–15. https://doi.org/10.1177/1463949115627895.

Moss, Peter, and Mathias Urban. 2019. 'The Organisation for Economic Co-operation and Development's International Early Learning Study: What's Going on'. *Contemporary Issues in Early Childhood* 20 (2): 207–12. https://doi.org/10.1177/1463949118803269.

———. 2020. 'The Organisation for Economic Co-operation and Development's International Early Learning and Child Well-Being Study: The Scores Are In!' *Contemporary Issues in Early Childhood* 21 (2): 165–71. https://doi.org/10.1177/1463949120929466.

Mundy, Karen. 1998. 'Educational Multilateralism and World (Dis)Order'. *Comparative Education Review* 42 (4): 448–78.

———. 2006. 'Education for All and the New Development Compact'. *International Review of Education* 52 (1/2): 23–48. https://doi.org/10.1007/s11159-005-5610-6.

———. 2007. 'Global Governance, Educational Change'. *Comparative Education* 43 (3): 339–57. https://doi.org/10.1080/03050060701556281.

Mundy, Karen, and Caroline Manion. 2015. 'The Education for All Initiative: History and Prospects Post-2015'. In *Education and International Development: An Introduction*, edited by Tristan McCowan and Elaine Unterhalter. London: Bloomsbury.

Mundy, Karen, and Antoni Verger. 2015. 'The World Bank and the Global Governance of Education in a Changing World Order'. *International Journal of Educational Development* 40: 9–18. https://doi.org/10.1016/j.ijedudev.2014.11.021.

Murphy, Craig. 1994. *International Organisation and Industrial Change: Global Governance since 1850*. Europe and the International Order. Cambridge: Polity.

National Research Council. 1995. *International Comparative Studies in Education: Descriptions of Selected Large-Scale Assessments and Case Studies*. Washington, D.C.: The National Academies Press. https://doi.org/10.17226/9174.

Neisser, Ulric. 1967. *Cognitive Psychology*. Century Psychology Series. London: Prentice-Hall.

Niemann, Dennis, and Kerstin Martens. 2018. 'Soft Governance by Hard Fact? The OECD as a Knowledge Broker in Education Policy'. *Global Social Policy* 18 (3): 267–83. https://doi.org/10.1177/1468018118794076.

Nóvoa, António. 2017. 'Comparative Education Disillusions: Politics and Knowledge'. CIES 2017 George F. Kneller Lecture presented at the Comparative and International Education Society Annual Conference, Atlanta.

———. 2018. 'Comparing Southern Europe: The Difference, the Public, and the Common'. *Comparative Education* 54 (4): 548–61. https://doi.org/10.1080/03050068.2018.1528781.

Nóvoa, Antonio, and Tali Yariv-Mashal. 2003. 'Comparative Research in Education: A Mode of Governance or a Historical Journey?' *Comparative Education* 39 (4): 423–38. https://doi.org/10.1080/0305006032000162002.

OECD. 1990. 'Acts of the Organisation Volume I'. OECD Publishing. EXT_OCDE_1_1990_Year 1992.pdf. https://oecd-learning.learning.aau.dk/search/results/documents.Document/?q=Acts+of+the+Organisation+1990.

———. 1995. *Governance in Transition: Public Management Reforms in OECD Countries.* Paris: OECD Publishing.

———. 1996a. 'Shaping the 21st Century: The Contribution of Development Co-operation'. Paris: OECD Publishing. https://www.oecd.org/dac/2508761.pdf.

———. 1996b. 'The Knowledge-Based Economy'. OECD Publishing. https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD%2896%29102&docLanguage=En.

———. 1999. 'Measuring Student Knowledge and Skills: A New Framework for Assessment'. Paris: OECD Publishing. https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33693997.pdf.

———. 2001. 'Background Paper'. Definition and Selection of Competencies (DeSeCo): Theoretical and Conceptual Foundations. Paris: OECD Publishing. http://www.oecd.org/education/skills-beyond-school/41529556.pdf

———. 2002. 'Strategy Paper'. Definition and Selection of Competencies (DeSeCo): Theoretical and Conceptual Foundations. Paris: OECD Publishing. https://www.deseco.ch/bfs/deseco/en/index/02.parsys.34116.downloadList.87902.DownloadFile.tmp/oecddesecostrategypaperdeelsaedcericd20029.pdf

———. 2003a. 'Rome Declaration on Harmonisation'. The High-Level Fora on Aid Effectiveness. Paris: OECD Publishing. https://www.oecd.org/dac/effectiveness/31451637.pdf.

———. 2003b. 'Summary of the Final Report: Key Competencies for a Successful Life and a Well-Functioning Society'. Paris: OECD Publishing. www.statistik.admin.ch/stat_ch/ber15/deseco/deseco_finalreport_summary.pdf.

———. 2004. 'Innovation in the Knowledge Economy: Implications for Education and Learning'. Paris: OECD Publishing. https://www.oecd-

ilibrary.org/education/innovation-in-the-knowledge-economy_9789264105621-en.

———. 2005a. 'Paris Declaration on Aid Effectiveness'. The High-Level Fora on Aid Effectiveness. Paris: OECD Publishing. https://doi.org/10.1787/9789264098084-en.

———. 2005b. 'The Definition and Selection of Key Competencies: Executive Summary'. Paris: OECD Publishing. https://www.deseco.ch/bfs/deseco/en/index/02.parsys.43469.downloadList.2296.DownloadFile.tmp/2005.dskcexecutivesummary.en.pdf

———. 2008. 'Accra Agenda for Action'. The High-Level Fora on Aid Effectiveness. Paris: OECD Publishing. https://dx.doi.org/10.1787/9789264098107-en.

———. 2011a. 'Better Policies for Development: Report on the DevGoals Exercise'. C/MIN(2011)9. Meeting of the OECD Council at Ministerial Level, 25-26 May 2011. Paris: OECD Publishing. https://www.oecd.org/mcm/48108568.pdf.

———. 2011b. 'Busan Partnership for Effectiveness Development Co-operation'. The High-Level Fora on Aid Effectiveness. Paris: OECD Publishing. https://www.oecd.org/dac/effectiveness/49650173.pdf.

———. 2011c. 'OECD 50th Anniversary Vision Statement'. C/MIN(2011)6. Meeting of the OECD Council at Ministerial Level, 25-26 May 2011. Paris: OECD Publishing. https://www.oecd.org/mcm/48064973.pdf

———. 2011d. 'Secretary-General's Strategic Orientations for 2011 and beyond'. Meeting of the OECD Council at Ministerial Level, 25-26 May 2011. Paris: OECD Publishing. https://www.oecd.org/mcm/48064885.pdf.

———. 2011e. 'The OECD at 50: Development Co-operation Past, Present and Future', 27–37. Paris: OECD Publishing. https://doi.org/10.1787/dcr-2011-7-en.

———. 2011f. 'Framework for an OECD Strategy on Development'. C/MIN(2011)8. Meeting of the OECD Council at Ministerial Level, 25-26 May 2011. Paris: OECD Publishing. https://www.oecd.org/governance/pcsd/Framework_OECD_Strategy_on_Development.pdf.

———. 2012a. 'Background Report for the OECD Strategy on Development'. C/MIN(2012)7. Meeting of the Council at Ministerial Level, 23-24 May 2012. Paris: OECD Publishing. https://www.oecd.org/gov/pcsd/Background%20Report_OECD_Strategy_on_Development.pdf.

———. 2012b. 'Framework for an OECD Strategy on Development'. Paris: OECD Publishing. https://www.oecd.org/development/48106820.pdf.

———. 2012c. 'Historical Summary of CERI's Main Activities'. Paris: OECD Publishing. http://www.oecd.org/education/ceri/50234306.pdf

———. 2012d. 'OECD Strategy on Development'. C/MIN(2012)6. Meeting of the Council at Ministerial Level, 23-24 May 2012. Paris: OECD Publishing. https://www.oecd.org/gov/pcsd/OECD_Strategy_on_Development.pdf.

———. 2013a. 'PISA 2012 Results (Volume IV): What Makes Schools Successful? Resources, Policies and Practices'. Paris: OECD Publishing. doi:10.1787/9789264201156-en.

———. 2013b. 'Progress Report on the Implementation of the OECD Strategy on Development'. C/MIN(2013)11. Meeting of the Council at Ministerial Level, 29-30 May 2013. Paris: OECD Publishing. https://www.oecd.org/gov/pcsd/Progress_Report_Implementation_OECD_Strategy_on_Development.pdf.

———. 2013c. 'The OECD's Contribution on Education to the Post-2015 Framework: PISA for Development'. 1. The OECD Post-2015 Reflection Series. Paris: OECD Publishing. https://www.oecd.org/dac/POST-2015%20paper%20education.pdf.

———. 2013d. 'Improving Learning Outcomes Worldwide: How PISA Can Help'. Paris: OECD Publishing. https://www.oecd.org/pisa/aboutpisa/pisa-for-development-brochure.pdf.

———. 2013e. 'PISA for Development Project Document (with Logical Framework)'. Paris: OECD Publishing. http://www.oecd.org/callsfortenders/ANNEX%20C.pdf.

———. 2014a. '2014 Report on the Implementation of the OECD Strategy on Development'. C/MIN(2014)14. Meeting of the OECD Council at Ministerial Level Paris, 6-7 May 2014. Paris: OECD Publishing. http://www.oecd.org/mcm/C-MIN%282014%2914-ENG.pdf.

———. 2014b. 'Call for Tender 100000990: PISA for Development Strand A and Strand B'. Paris: OECD Publishing. http://www.oecd.org/callsfortenders/2014%2008%2004%20PISA%20for%20DEV%20CFT%20Strands%20AB.pdf.

———. 2014c. 'Lessons from PISA for Korea: Strong Performers and Successful Reformers in Education'. Paris: OECD Publishing. https://doi.org/10.1787/9789264190672-en.

———. 2014d. 'Looking Ahead to Global Development beyond 2015: Lessons Learnt from the Initial Implementation Phase of the OECD Strategy on Development'. C/MIN(2014)13. Meeting of the OECD Council at Ministerial Level Paris, 6-7 May 2014. Paris: OECD Publishing. http://www.oecd.org/mcm/C-MIN%282014%2914-ENG.pdf.

———. 2015. 'Call for Tenders 100001420: International Early Learning Study'. Paris: OECD Publishing. http://www.oecd.org/callsfortenders/CfT%20100001420%20International%20Early%20Learning%20Study.pdf.

———. 2016a. 'Capacity Building Plan: Cambodia'. PISA for Development. Paris: OECD Publishing. https://www.oecd.org/pisa/aboutpisa/Cambodia-Capacity-Building-Plan.pdf.

———. 2016b. 'Making Education Count for Development: Data Collection and Availability in Six PISA for Development Countries'. PISA. Paris: OECD Publishing. https://dx.doi.org/10.1787/9789264255449-1-en.

———. 2016c. 'PISA for Development Brochure'. Paris: OECD Publishing. http://www.oecd.org/pisa/pisa-for-development/pisa-for-development-background.htm.

———. 2016d. 'PISA for Development: Enhancement of the Cognitive Instruments'. Paris: OECD Publishing. https://www.oecd.org/pisa/aboutpisa/5%20-%20Enhancements%20to%20cognitive%20instruments.pdf.

———. 2016e. 'PISA for Development: Benefits for Participating Countries'. Paris: OECD Publishing. https://www.oecd.org/pisa/aboutpisa/PISA-for-Development-Benefits-for-participating-countries-PISA-D-Brief2.pdf.

———. 2016f. 'How Is PISA for Development Being Implemented?' Paris: OECD Publishing. https://www.oecd.org/pisa/aboutpisa/4%20-%20How%20is%20PISA-D%20implemented.pdf.

———. 2017a. 'Development Co-operation Report 2017: Data for Development'. Text. Development Co-operation Report. Paris: OECD Publishing. https://doi.org/10.1787/dcr-2017-en.

———. 2017b. 'Sampling and Survey Operations for the PISA for Development Out-of-School Assessment Field Trial'. Paris: OECD Publishing. https://www.oecd.org/pisa/pisa-for-development/15-Out-of-school-sample-and-operations.pdf.

———. 2017c. 'PISA 2015 Context Questionnaires Framework'. PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving. Paris: OECD Publishing. https://doi.org/10.1787/9789264281820-7-en.

———. 2018a. 'OECD Learning Framework 2030'. Position paper. OECD Future of Education and Skills 2030. Paris: OECD Publishing. https://www.oecd.org/education/2030-project/about/documents/E2030%20Position%20Paper%20(05.04.2018).pdf.

———. 2018b. 'PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science'. Paris: OECD Publishing. https://www.oecd-ilibrary.org/education/pisa-for-development-assessment-and-analytical-framework_9789264305274-en.

———. 2018c. 'PISA for Development: Results in Focus'. Paris: OECD Publishing. https://www.oecd-ilibrary.org/education/pisa-for-development_c094b186-en.

———. 2018d. 'The Power and Promise of Early Learning'. OECD Education Working Paper No. 186. Paris: OECD Publishing. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/WKP%20%282018%2922&docLanguage=En.

———. 2018e. 'Strengthening Analysis and Reporting in PISA for Development Countries'. Paris: OECD Publishing. https://www.oecd.org/pisa/pisa-for-development/25_Strengthening_analysis_reporting_in%20PISA_D_countries.pdf.

———. 2018f. 'Peer-to-Peer Learning to Strengthen Dissemination of PISA for Development Results'. Paris: OECD Publishing.

https://www.oecd.org/pisa/pisa-for-development/27_Peer_learning_results_dissemination.pdf.

———. 2019a. 'Meeting Expectations? Reflections on PISA for Development from Participating Countries'. Paris: OECD Publishing. https://www.oecd.org/pisa/pisa-for-development/Brief_29_Meeting_expectations.pdf.

———. 2019b. 'OECD Learning Compass 2030'. A Series of Concept Notes. OECD Future of Education and Skills 2030. Paris: OECD Publishing. https://www.oecd.org/education/2030-project/about/documents/E2030%20Position%20Paper%20(05.04.2018).pdf.

———. 2019c. 'PISA for Development Project Completion Report'. PISA for Development International Seminar, 25 September 2019 London. Paris: OECD Publishing. http://www.oecd.org/pisa/pisa-for-development/PISA%20for%20Development%20Project%20Completion%20Report.pdf.

———. 2019d. 'Summary Record of the PISA for Development International Seminar'. Paris: OECD Publishing. http://www.oecd.org/pisa/pisa-for-development/Summary_record_PISA-D_International_Seminar.pdf.

———. 2020. 'PISA for Development: Out-of-School Assessment: Results in Focus'. Paris: OECD Publishing. https://www.oecd-ilibrary.org/education/pisa-for-development-out-of-school-assessment_491fb74a-en.

Overbeek, Henk, Klaus Dingwerth, Philipp Pattberg, and Daniel Compagnon. 2010. 'Forum: Global Governance: Decline or Maturation of an Academic Concept?' *International Studies Review* 12 (4): 696–719.

Ozga, Jennifer, Peter Dahler-Larsen, Christina Segerholm, and Hannu Simola. 2011. *Fabricating Quality in Education Data and Governance in Europe*. 1st ed. London: Routledge.

Ozga, Jenny. 2012. 'Assessing PISA'. *European Educational Research Journal* 11 (2): 166–71. https://doi.org/10.2304/eerj.2012.11.2.166.

Papadopoulos, George S. 1994. *Education 1960-1990: The OECD Perspective*. OECD Historical Series. Paris: OECD Publishing.

———. 2011. 'The OECD Approach to Education in Retrospect: 1960-1990'. *European Journal of Education* 46 (1): 85–86.

Patel, Dev, and Justin Sandefur. 2020. 'A Rosetta Stone for Human Capital'. Working Paper 550. Washington, D.C.: Center for Global Development. https://www.cgdev.org/sites/default/files/rosetta-stone-human-capital.pdf.

Pettersson, Daniel, and Thomas S. Popkewitz. 2019. 'A Chimera of Quantifications and Comparisons: The Changing of Educational "Expertise"'. In *New Practices of Comparison, Quantification and Expertise in Education: Conducting Empirically Based Research*, edited by Christina Mølstad and Daniel Pettersson, 18–36. Oxon: Routledge.

Phillips, Nelson, and John L. Brown. 1993. 'Analysing Communication in and around Organisations: A Critical Hermeneutic Approach'. *Academy of Management Journal* 36 (6): 1547–76. https://doi.org/10.2307/256821.

Phillips, Nelson, and Cynthia Hardy. 2002. 'What Is Discourse Analysis?' In *Discourse Analysis*, 2–17. Thousand Oaks: SAGE Publications. https://doi.org/10.4135/9781412983921.

Pizmony-Levy, Oren. 2013. 'Testing for All: The Emergence and Development of International Assessment of Student Achievement, 1958–2012'. ProQuest Dissertations Publishing. https://search.proquest.com/docview/1461405340?pq-origsite=summon.

Potter, Jonathan. 1996. *Representing Reality: Discourse, Rhetoric and Social Construction*. London: SAGE Publications. http://dx.doi.org/10.4135/9781446222119.

Pritchett, Lant. 2012. 'The First PISA Results for India: The End of the Beginning'. 5 January 2012. https://blog.theleapjournal.org/2012/01/first-pisa-results-for-india-end-of.html.

Psacharopoulos, George. 1994. 'Returns to Investment in Education: A Global Update'. *World Development*, World Development, 22 (9): 1325–43. https://doi.org/10.1016/0305-750x(94)90007-8.

Psacharopoulos, George, and Harry A. Patrinos. 2004. 'Returns to Investment in Education: A Further Update'. *Education Economics*, Education Economics, 12 (2): 111–34. https://doi.org/10.1080/0964529042000239140.

Punch, Keith. 2014. *Introduction to Research Methods in Education*. 2nd ed. Los Angeles: SAGE Publications.

Ramirez, Francisco O., Xiaowei Luo, Evan Schofer, and John W. Meyer. 2006. 'Student Achievement and National Economic Growth'. *American Journal of Education* 113 (1): 1–29. https://doi.org/10.1086/506492.

Ramirez, Francisco O., Evan Schofer, and John W. Meyer. 2018. 'International Tests, National Assessments, and Educational Development (1970–2012)'. *Comparative Education Review* 62 (3): 344–64. https://doi.org/10.1086/698326.

Rappleye, Jeremy, and Hikaru Komatsu. 2019. 'Is Knowledge Capital Theory Degenerate? PIAAC, PISA, and Economic Growth'. *Compare: A Journal of Comparative and International Education*, 1–19. https://doi.org/10.1080/03057925.2019.1612233.

Rappleye, Jeremy, Hikaru Komatsu, Yukiko Uchida, Kuba Krys, and Hazel Markus. 2020. '"Better Policies for Better Lives"?: Constructive Critique of the OECD's (Mis)Measure of Student Well-Being'. *Journal of Education Policy* 35 (2): 258–82. https://doi.org/10.1080/02680939.2019.1576923.

Resnik, Julia. 2006. 'International Organisations, the "Education–Economic Growth" Black Box, and the Development of World Education Culture'. *Comparative Education Review* 50 (2): 173–95. https://doi.org/10.1086/500692.

Ricoeur, Paul. 1971. 'The Model of the Text: Meaningful Action Considered as a Text'. *Social Research* 38 (3): 529–62.

Riessman, Catherine K. 2002. 'Narrative Analysis'. In *The Qualitative Researcher's Companion*, edited by Michael A. Huberman and Matthew B. Miles, 216–70. Thousand Oaks California: SAGE Publications. https://doi.org/10.4135/9781412986274.n10.

Rinne, Risto, Johanna Kallo, and Sanna Hokka. 2004. 'Too Eager to Comply? OECD Education Policies and the Finnish Response'. *European Educational Research Journal* 3 (2): 454–85. https://doi.org/10.2304/eerj.2004.3.2.3.

Rittberger, Volker, and Bernhard Zangl. 2006. *International Organisation: Polity, Politics and Policies*. Basingstoke: Palgrave Macmillan.

Rizvi, Fazal, and Bob Lingard. 2006. 'Education Globalisation and Social Change'. In *Globalization and the Changing Nature of the OECD's Educational Work*. Oxford University Press. http://minerva-access.unimelb.edu.au/handle/11343/31469.

———. 2009. 'The OECD and Global Shifts in Education Policy'. In *International Handbook of Comparative Education*, edited by Robert Cowen and Andreas M. Kazamias, 437–53. Springer International Handbooks of Education. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-6403-6_28.

Roberts, Peter. 2007. 'Neoliberalism, Performativity and Research'. *International Review of Education* 53 (4): 349–65. https://doi.org/10.1007/s11159-007-9049-9.

Robertson, Susan L. 2014. 'From Thin to Thickening Global Governance of Education'. *NORRAG News*, 2014.

Rosenau, James N. 1992. 'Governance, Order, and Change in World Politics'. In *Governance without Government: Order and Change in World Politics*, edited by Ernst-Otto Czempiel and James N. Rosenau, 1–29. Cambridge Studies in International Relations. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511521775.003.

Rosenau, James N., and Ernst-Otto Czempiel. 1992. *Governance without Government: Order and Change in World Politics*. Cambridge Studies in International Relations; No. 20. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511521775.

Rostow, Walt W. 1960. *The Stages of Economic Growth: A Non-Communist Manifesto*. 1st ed. Cambridge: Cambridge University Press.

Rottenburg, Richard, Sally E. Merry, Sung-Joon Park, and Johanna Mugler, eds. 2015. *The World of Indicators: The Making of Governmental Knowledge through Quantification*. Cambridge Studies in Law and Society. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781316091265.

Rubenson, Kjell. 2015. 'Framing the Adult Learning and Education Policy Discourse: The Role of the Organization for Economic Co-operation and Development'. In *Global Perspectives on Adult Education and Learning Policy*, edited by Marcella Milana and Tom Nesbit, 179–93. Palgrave Studies in Global Citizenship Education and Democracy. London: Palgrave Macmillan. https://doi.org/10.1057/9781137388254_13.

Ruggie, John G. 1998. 'What Makes the World Hang Together? Neo-Utilitarianism and the Social Constructivist Challenge'. *International Organisation* 52 (4): 855–85. https://doi.org/10.1162/002081898550770.

Sadler, Michael. 1900. *How Far Can We Learn Anything of Practical Value from the Study of Foreign Systems of Education?* Guildford: Surrey Advertiser.

Said, Edward W. 1978. *Orientalism*. London: Routledge and Kegan Paul.

Salajan, Florin D., and Tavis D. Jules. 2020. 'Introduction: The Educational Intelligent Economy, Educational Intelligence, and Big Data'. In *The Educational Intelligent Economy: Big Data, Artificial Intelligence, Machine Learning and the Internet of Things in Education*, edited by Tavis D. Jules and Florin D. Salajan, 1st ed., 38:1–11. International Perspectives on Education and Society. Bingley: Emerald Publishing Limited. https://doi.org/10.1108/S1479-367920190000038001.

Scherman, Vanessa, Roel J. Bosker, and Sarah J. Howie. 2017. *Monitoring the Quality of Education in Schools: Examples of Feedback into Systems from Developed and Emerging Economies*. Rotterdam, The Netherlands: Sense Publishers.

Schleicher, Andreas. 2013a. 'PISA for Development Overview Presentation'. Presented at the PISA for Development Initial Technical Meeting, Paris. https://www.oecd.org/pisa/aboutpisa/pisa-for-development-meetings.htm.

———. 2013b. 'Big Data and PISA'. OECD Education and Skills Today. 26 June 2013. https://oecdedutoday.com/big-data-and-pisa/.

———. 2018. 'World Class: How to Build a 21st-Century School System'. Strong Performers and Successful Reformers in Education. Paris: OECD Publishing. http://dx.doi.org/10.1787/4789264300002-en.

———. 2019. *Principal Drivers and Rationale for Enhancing PISA in the Context of Sustainable Development Goals (SDGs)*. Presented at the PISA for Development International Seminar. London. http://www.oecd.org/pisa/pisa-for-development/pisafordevelopmentinternationalseminar.htm.

———. 2020. Helping Students Navigate in an Interconnected World – What to Expect from 2018 (Volume VI) Online Webinar. https://fb.watch/2U34Sv0HOU/.

Schleicher, Andreas, and Claudia Costin. 2015. 'The Challenges of Widening Participation in PISA'. *World Bank Blogs* (blog). 27 November 2015. https://blogs.worldbank.org/education/challenges-widening-participation-pisa.

Schriewer, Jürgen. 2012. 'EDITORIAL: Meaning Constellations in the World Society'. *Comparative Education* 48 (4): 411–22. https://doi.org/10.1080/03050068.2012.737233.

Schriewer, Jürgen, and Carlos Martinez. 2004. 'Constructions of Internationality in Education'. In *The Global Politics of Educational Borrowing and Lending*, edited by Gita Steiner-Khamsi, 29–53. New York: Teachers College Press.

Schultz, Theodore W. 1961. 'Investment in Human Capital'. *The American Economic Review* 51 (1): 1–17. http://www.jstor.org/stable/1818907.

———. 1971. *Investment in Human Capital: The Role of Education and of Research*. New York: Free Press.

Schutz, Alfred. 1962. *Collected Papers I: The Problem of Social Reality*. Edited by Maurice Natanson. The Hague: Martinus Nijhoff.

Schwartz-Shea, Peregrine, and Dvora Yanow. 2012. *Interpretive Research Design: Concepts and Processes*. Routledge Series on Interpretive Methods. New York: Routledge. http://dx.doi.org/10.4324/9780203854907.

Sellar, Sam, and Bob Lingard. 2014. 'The OECD and the Expansion of PISA: New Global Modes of Governance in Education'. *British Educational Research Journal* 40 (6): 917–36. https://doi.org/10.1002/berj.3120.

Shamatov, Duishon, and Keneshbek Sainazarov. 2010. 'The Impact of Standardised Testing on Education Quality in Kyrgyzstan: The Case of the Program for International Student Assessment (PISA) 2006'. In *The Impact of International Achievement Studies on National Education Policymaking*, edited by Alexander W. Wiseman, 13:145–79. International Perspectives on Education and Society. Emerald Group Publishing Limited. https://doi.org/10.1108/S1479-3679(2010)0000013009.

Silova, Iveta, and Euan Auld. 2019. '"Banal" Imperialism and International Education Performance Data: Breaching the World Order by Numbers'. *International Journal for the Historiography of Education* 9 (2): 269–274 https://www.klinkhardt.de/newsite/media/20191022_9783781523388%20IJHE-2%20Inh%20Editorial.pdf

Silova, Iveta, and William C. Brehm. 2015. 'From Myths to Models: The (Re)Production of World Culture in Comparative Education'. *Globalisation, Societies and Education* 13 (1): 8–33. https://doi.org/10.1080/14767724.2014.967483.

Silva, Gabriela T. 2019. 'International Cooperation from the Perspective of INEP Agents: The OECD and Brazilian Public Education, 1996–2006'. In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 109–32. Global Histories of Education. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-33799-5.

Smith, Katherine E. 2006. 'Problematising Power Relations in "Elite" Interviews'. *Geoforum* 37 (4): 643–53. https://doi.org/10.1016/j.geoforum.2005.11.002.

Solheim, Erik. 2015. 'PISA for Development'. *OECD Education and Skills Today* (blog). 30 March 2015. https://oecdedutoday.com/pisa-for-development/.

Spivak, Gayatri C. 1988. 'Can the Subaltern Speak?' In *Marxism and the Interpretation of Culture*, edited by Cary Nelson and Lawrence Grossberg, 271–316. Urbana: University of Illinois Press.

———. 1999. *A Critique of Postcolonial Reason: Toward a History of the Vanishing Present*. Massachusetts: Harvard University Press.

Steiner-Khamsi, Gita. 2004. *The Global Politics of Educational Borrowing and Lending*. New York: Teachers College Press.

———. 2017. 'Focusing on the Local to Understand Why the Global Resonates and How Governments Appropriate ILSAs for National Agenda Setting'. *Compare: A Journal of Comparative and International Education* 47 (3): 443–46. https://doi.org/10.1080/03057925.2017.1301399.

Sternberg, Robert J. 1981. 'Testing and Cognitive Psychology'. *The American Psychologist* 36 (10): 1181–89. https://doi.org/10.1037/0003-066X.36.10.1181.

Stromquist, Nelly P. 2016. 'Using Regression Analysis to Predict Countries' Economic Growth: Illusion and Fact in Education Policy'. *Real-World Economics Review*, no. 76: 65–74. http://www.paecon.net/PAEReview/issue76/Stromquist76.pdf

Suchman, Mark C. 1995. 'Managing Legitimacy: Strategic and Institutional Approaches'. *The Academy of Management Review* 20 (3): 571–610. https://doi.org/10.2307/258788.

Sugrue, Brenda. 1995. 'A Theory-Based Framework for Assessing Domain-Specific Problem-Solving Ability'. *Educational Measurement: Issues and Practice* 14 (3): 29–35. https://doi.org/10.1111/j.1745-3992.1995.tb00865.x.

Sullivan, Harry S. 1954. *The Psychiatric Interview*. New York: W.W. Norton.

Takayama, Keita. 2008. 'The Politics of International League Tables: PISA in Japan's Achievement Crisis Debate'. *Comparative Education* 44 (4): 387–407. https://doi.org/10.1080/03050060802481413.

Taylor, Charles. 1971. 'Interpretation and the Sciences of Man'. *The Review of Metaphysics* 25 (1): 3–51. http://www.jstor.org/stable/20125928.

The Global Partnership for Education. 2018. 'Global Partnership for Education Results Report 2018'. Washington, D.C.: The Global Partnership for Education. https://www.globalpartnership.org/sites/default/files/2018-06-gpe-results-report-2018-web4.pdf.

Thuesen, Frederik. 2011. 'Navigating between Dialogue and Confrontation: Phronesis and Emotions in Interviewing Elites on Ethnic Discrimination'. *Qualitative Inquiry* 17 (7): 613–22. https://doi.org/10.1177/1077800411413998.

Tikly, Leon. 2004. 'Education and the New Imperialism'. *Comparative Education* 40 (2): 173–98. https://doi.org/10.1080/0305006042000231347.

———. 2017. 'The Future of Education for All as a Global Regime of Educational Governance'. *Comparative Education Review* 61 (1): 22–57. https://doi.org/10.1086/689700.

Tröhler, Daniel. 2011. 'The Global Language on Education Policy and Prospects of Education Research'. In *Education Systems in Historical, Cultural, and Sociological Perspectives*, edited by Daniel Tröhler and Ragnhild Barbu, 55–73. The Future of Education Research. Rotterdam: Sense Publishers. https://doi.org/10.1007/978-94-6091-827-8_5.

———. 2013. 'The OECD and Cold War Culture: Thinking Historically about PISA'. In *PISA, Power, and Policy: The Emergence of Global Educational Governance*, edited by Aaron Benavot and Heniz-Dieter Meyer, 141–61. Oxford: Symposium Books.

UIS Global Education Monitoring (GEM) Report Team. 2016. 'SDG4 Is a Universal Agenda – and That Includes High Income Countries'. *World Education Blog* (blog). 16 September 2016. https://gemreportunesco.wordpress.com/2016/09/16/sdg4-is-a-universal-agenda-and-that-includes-high-income-countries/.

Ulich, Robert, ed. 1964. *Education and the Idea of Mankind*. Chicago: The University of Chicago Press.

UNESCO. 2015. 'Education 2030 Incheon Declaration and Framework for Action'. Paris: UNESCO. http://www.uis.unesco.org/Education/Documents/incheon-framework-for-action-en.pdf.

UNESCO Institute for Statistics. 2016a. 'Sustainable Development Data Digest: Laying the Foundation to Measure Sustainable Development Goal 4'. Montreal: UNESCO Institute for Statistics. http://uis.unesco.org/sites/default/files/documents/laying-the-foundation-to-measure-sdg4-sustainable-development-data-digest-2016-en.pdf.

———. 2016b. 'The Cost of Not Assessing Learning Outcomes'. Montreal: UNESCO Institute for Statistics. http://uis.unesco.org/sites/default/files/documents/the-cost-of-not-assessing-learning-outcomes-2016-en_0.pdf.

UNESCO International Institute for Educational Planning (2016). An Insider's View of the Programme for International Student Assessment (PISA) for Development (PISA-D) initiative Interview by UNESCO International Institute for Educational Planning Staff Member on 19 February 2016. https://www.youtube.com/watch?v=qLiWkz4eaVs.

UNICEF. 2018. 'Progress for Every Child in the SDG Era'. New York: UNICEF. https://data.unicef.org/resources/progress-for-every-child-in-the-sdg-era-2020/.

UNICEF, and UNESCO Institute for Statistics. 2010. 'All in School: Global Initiative in out-of-School Children'. http://allinschool.org/http://www.uis.unesco.org/Education/Documents/oosci-global-report-en.pdf.

Unterhalter, Elaine. 2009. 'What Is Equity in Education? Reflections from the Capability Approach'. *Studies in Philosophy and Education* 28 (5): 415–24. https://doi.org/10.1007/s11217-009-9125-7.

———. 2015. 'Education and International Development: A History of the Field'. In *Education and International Development: An Introduction*, edited by Tristan McCowan and Elaine Unterhalter, 13–30. London: Bloomsbury.

———. 2019. 'The Many Meanings of Quality Education: Politics of Targets and Indicators in SDG4'. *Global Policy* 10 (S1): 39–51. https://doi.org/10.1111/1758-5899.12591.

Urry, John. 2003. 'Social Networks, Travel and Talk'. *The British Journal of Sociology* 54 (2): 155–75. https://doi.org/10.1080/0007131032000080186.

———. 2004. 'Small Worlds and the New "Social Physics"'. *Global Networks* 4 (2): 109–30. https://doi.org/10.1111/j.1471-0374.2004.00083.x.

Valverde, Gilbert A. 2014. 'Educational Quality: Global Politics, Comparative Inquiry, and Opportunities to Learn'. *Comparative Education Review* 58 (4): 575–89. https://doi.org/10.1086/678038.

Vaughan, Rosie P. 2015. 'Research Paradigms'. In *Education and International Development: An Introduction*, edited by Tristan McCowan and Elaine Unterhalter, 79–90. London: Bloomsbury.

Verger, Antoni, Mario Novelli, and Hülya K. Altinyelken, eds. 2012. *Global Education Policy and International Development: New Agendas, Issues and Policies*. 1st ed. London: Bloomsbury Academic. https://doi.org/10.5040/9781472544575.

Viotti, Paul R., and Mark V. Kauppi. 1987. *International Relations Theory: Realism, Pluralism, Globalism*. 1st ed. New York: Macmillan.

Vygotskii, Lev S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Edited by Michael Cole, Vera John-Steiner, Sylvia Scribner, and Ellen Souberman. Cambridge: Harvard University Press. http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=nlebk&AN=575543&site=ehost-live&scope=site.

Wagner, Daniel A. 2010. 'Quality of Education, Comparability, and Assessment Choice in Developing Countries'. *Compare: A Journal of Comparative and International Education* 40 (6): 741–60. https://doi.org/10.1080/03057925.2010.523231.

———. 2011. *Smaller, Quicker, Cheaper: Improving Learning Assessments for Developing Countries*. Paris: International Institute for Educational Planning.

———. 2018. *Learning as Development: Rethinking International Education in a Changing World*. New York: Routledge. https://doi.org/10.4324/9780203115305.

Wagner, Daniel A., Sharon Wolf, and Robert F. Boruch, eds. 2018. *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries*. Paris: UNESCO IIEP.

Wagner, Daniel, and Nathan Castillo. 2014. 'Learning at the Bottom of the Pyramid: Constraints, Comparability and Policy in Developing Countries'. *PROSPECTS* 44 (4): 627–38. https://doi.org/10.1007/s11125-014-9328-8.

Walby, Kevin. 2010. 'Interviews as Encounters: Issues of Sexuality and Reflexivity When Men Interview Men about Commercial Same Sex Relations'. *Qualitative Research* 10 (6): 639–57. https://doi.org/10.1177/1468794110380525.

Waldow, Florian. 2012. 'Standardisation and Legitimacy: Two Central Concepts in Research on Educational Borrowing and Lending'. In *World Yearbook of Education 2012*, edited by Gita Steiner-Khamsi and Florian Waldow, 1st ed., 411–27. London: Routledge. https://doi.org/10.4324/9780203137628-33.

Waldow, Florian, Keita Takayama, and Youl-Kwan Sung. 2014. 'Rethinking the Pattern of External Policy Referencing: Media Discourses over the "Asian Tigers'" PISA Success in Australia, Germany and South Korea'. *Comparative Education* 50 (3): 302–21. https://doi.org/10.1080/03050068.2013.860704.

Walsh, David. 1972. 'Sociology and the Social World'. In *New Directions in Sociological Theory*, by Paul Filmer, Michael Phillipson, David Silverman, and David Walsh, 246. London: Collier-Macmillan.

Walsham, Geoffrey. 1993. *Interpreting Information Systems in Organisations*. John Wiley Series in Information Systems. Chichester: John Wiley.

Ward, Michael. 2016. 'PISA for Development'. Phnom Penh, February 29. http://www.unescobkk.org/fileadmin/user_upload/epr/Quality/NEQMAP/1-1.PISA_for_Development_Michael_Ward.pdf.

———. 2017. *Welcome Speech*. The 5th National Project Managers Meeting of PISA for Development. Siem Reap, Cambodia.

Ward, Michael, and Pablo Zoido. 2015. 'PISA for Development'. In *Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik* 38 (4):21–25. Münster. https://doi.org/10.1787/c094b186-en.

Watson, John B. 1930. *Behaviorism*. Rev. Behaviorism, Rev. Ed. Chicago: University of Chicago Press.

Webb, Eugene, Donald Campbell, Richard Schwartz, and Lee Sechrest, eds. 2000. *Unobtrusive Measures*. Rev. Thousand Oaks: SAGE Publications. https://doi.org/10.4135/9781452243443.

Weber, Max. 1947. *Max Weber: The Theory of Social and Economic Organisation*. Translated by Talcott Pasons and A. M. Henderson. Illinois: Free Press.

Williamson, Ben. 2016. 'Digital Education Governance: An Introduction'. *European Educational Research Journal* 15 (1): 3–13. https://doi.org/10.1177/1474904115616630.

———. 2017. *Big Data in Education: The Digital Future of Learning, Policy and Practice*. 1st ed. Los Angeles: SAGE Publications. http://dx.doi.org/10.4135/9781529714920.

Willms, Douglas J. 2003. 'A Sense of Belonging and Participation - Results from PISA 2000'. Student Engagement at School. Paris: OECD Publishing. https://doi.org/10.1787/9789264018938-en.

———. 2006. 'Learning Divides: Ten Policy Questions about the Performance and Equity of Schools and Schooling Systems'. UIS Working Paper No.5. Montreal: UNESCO Institute for Statistics. http://uis.unesco.org/sites/default/files/documents/learning-divides-ten-policy-questions-about-the-performance-and-equity-of-schools-and-schooling-systems-06-en_0.pdf.

———. 2011. Student Engagement: A Leadership Priority – An Interview with J. Douglas Willms *In Conversation*. http://www.edu.gov.on.ca/eng/policyfunding/leadership/summer2011.pdf.

———. 2018a. 'Educational Prosperity: An Assessment Strategy for Supporting Student Learning in Low-Income Countries'. In *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries*, edited by Daniel A. Wagner, Sharon Wolf, and Robert F. Boruch, 97–113. Paris: UNESCO-IIEP.

———. 2018b. 'Learning Divides: Using Data to Inform Educational Policy'. Information Paper 54. Montreal: UNESCO Institute for Statistics. http://uis.unesco.org/sites/default/files/documents/ip54-learning-divides-using-data-inform-educational-policy.pdf.

Willms, Douglas J., and Lucia Tramonte. 2015. 'Towards the Development of Contextual Questionnaires for the PISA for Development Study'. OECD Education Working Papers. Paris: OECD Publishing. http://dx.doi.org/10.1787/5js1kv8crsjf-en.

Wiseman, Alexander W., and Calley S. Taylor. 2017. *The Impact of the OECD on Education Worldwide*. International Perspectives on Education and Society. Bingley: Emerald Publishing Limited. http://www.emeraldinsight.com/doi/book/10.1108/S1479-3679201731.

Wolming, Simon, and Christina Wikström. 2010. 'The Concept of Validity in Theory and Practice'. *Assessment in Education: Principles, Policy & Practice* 17 (2): 117–32. https://doi.org/10.1080/09695941003693856.

World Bank. 2018. 'World Bank 2018 Development Report: Learning to Realise Education's Promise'. Washington, D.C.: World Bank. doi:10.1596/978-1-4648-1096-1.

———. 2019a. 'Ending Learning Poverty: What It Will Take'. Washington, D.C.: World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/32553/142659.pdf?sequence=7.

———. 2019b. 'World Bank, UNESCO Institute for Statistics Join Forces to Help Countries Measure Student Learning'. 8 July 2019. https://www.worldbank.org/en/topic/education/brief/world-bank-unesco-institute-for-statistics-join-forces-to-help-countries-measure-student-learning.

Wright, Georg H. von. 1971. *Explanation and Understanding*. International Library of Philosophy and Scientific Method. Ithaca: Cornell University Press.

Wulff, Antonia. 2019. 'The Ideological Battle over SDG 4'. Civil Society Report. Spotlight on Sustainable Development 2019. https://www.2030spotlight.org/sites/default/files/spot2019/Spotlight_Innenteil_2019_web_sdg4.pdf.

Yanow, Dvora. 2009. 'Dear Author, Dear Reader: The Third Hermeneutic in Writing and Reviewing Ethnography'. In *Political Ethnography: What Immersion Contributes to the Study of Power*, edited by Edward Schatz, 275–302. Chicago: University of Chicago Press. http://UCL.eblib.com/patron/FullRecord.aspx?p=3038320.

Yanow, Dvora, and Peregrine Schwartz-Shea, eds. 2015. 'Judging Quality: Evaluative Criteria and Epistemic Communities'. In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, 2nd ed., 552. New York: Routledge. https://doi.org/10.4324/9781315703275-8.

Ydesen, Christian. 2019. 'Introduction: What Can We Learn about Global Education from Historical and Global Policy Studies of the OECD?' *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 1–14. Global Histories of Education. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-33799-5.

Ydesen, Christian, and Sotiria Grek. 2020. 'Securing Organisational Survival: A Historical Inquiry into the OECD's Work in Education during the 1960s'. *Paedagogica Historica* 56 (3): 412–27. https://doi.org/10.1080/00309230.2019.1604774.

Ydesen, Christian, Alison L. Milner, Tali R. Aderet-German, Ezequiel G. Caride, Youjin Ruan, Ida S. Hansen, and Felicitas Acosta. 2020. 'Inclusion in Testing Times: Implications for Citizenship and Participation', August. https://vbn.aau.dk/en/publications/inclusion-in-testing-times-implications-for-citizenship-and-parti.

Zabusky, Stacia. 2007. 'Ethnography in/of Transnational Processes: Following Gyres in the Worlds of Big Science and European Integration'. In *Ethnographic Fieldwork: An Anthropological Reader*, edited by Antonius C. G. M. Robben and Jeffrey A. Sluka, Anthologies in Social and Cultural Anthropology. Malden: Blackwell.

Zacher, Mark W. 1979. *International Conflicts and Collective Security, 1946-77: United Nations, Organisation of American States, Organisation of African Unity and Arab League*. New York: Praeger.

Zuckerman, Harriet. 1972. 'Interviewing an Ultra-Elite'. *Public Opinion Quarterly* 36 (2): 159–75. https://doi.org/10.1086/267989.