Piloting PISA for Development to success: An analysis of its findings, framework and recommendations

Abstract

In 2018 the OECD published the findings of its PISA for Development (PISA-D) pilot project which was undertaken to make the regular PISA framework more accessible and relevant to low- and middle-income nations. This would encourage such nations to join PISA as part of the OECD's Learning Framework 2030 and provide them with 'contextualised' policy recommendations. In 2019 the OECD declared the project a success. We analyse and compare the PISA-D reports as well as its portrayal as a success. We suggest that, whilst PISA-D clearly made technical adjustments relating to the longstanding challenges which face low-income nations engaging in comparative assessments it replicates rather than addresses those challenges. Drawing on literature on organisational legitimacy and the politics of expertise, we interpret the PISA-D pilot as a political strategy primarily deployed to legitimate and extend rather than evaluate the project.

Keywords: PISA; OECD Learning Framework 2030; Sustainable Development Goals; low- and middle-income countries; organisational legitimacy

Introduction

> Now everybody supports PISA for Development because it's been so successful... We have now instruments that can extend throughout the world. We have proven that.
>
> Andreas Schleicher, PISA-D International Seminar[1], 2019

The Organisation for Economic Co-operation and Development's (OECD) Programme for International Student Assessment (PISA) assesses triennially the literacy and numeracy skills of 15-year olds across countries. This paper focuses on PISA for Development (hereafter PISA-D), undertaken by the OECD from 2013 to create a version of PISA that could accommodate low- and middle-income nations. In late 2018, the OECD released the report on the results of PISA-D, subsequently the seven pilot nations (i.e. Cambodia, Ecuador, Senegal, Zambia, Honduras, Guatemala and Paraguay[2]) published their country specific reports and in 2019, they declared PISA-D a success. Although these were not accompanied by the extensive media coverage that characterises the release of the OECD's regular assessment reports (see Grey and Morris 2018), the ambitions behind the initiative are significant.

Positioned as a 'pilot project' (OECD 2013, 2014a), PISA-D's stated objective was to adapt and enhance the relevance of the PISA testing instruments and contextual questionnaires for less affluent nations, while maintaining the comparability of the findings with the PISA standard (OECD 2018a). If successful, the initiative would support its strategic ambitions in the global governance of education (Li and Auld 2020). Currently PISA covers over 70 mainly affluent nations. The OECD's goal is to broaden PISA's coverage to 170 nations by 2030, so that the results from the 2021 PISA cycle would provide a consistent

---

1 The PISA-D International Seminar was held in London on 25 September 2019. It launched the PISA-D reports and discussed the outputs and findings. The Seminar was attended mainly by the OECD staff, contractors to the project and five representatives from participating nations.

2 Bhutan was excluded as the data was incomplete. Panama only conducted the out-of-school survey.

measure of progress towards the UN's Sustainable Development Goals for education (SDG 4). Recognising the potential impact of PISA-D on future aid disbursement and education governance post-2015 in low-income nations, and the lack of independent research systematically interrogating its outcomes, we analyse the project's findings, framework and recommendations.

The analysis complements research detailing the arrival of PISA-D in pilot nations (e.g. Auld et al. 2019) and research into the pilot process itself (Gorur et al. 2019; Addey and Gorur 2020). Specifically, we take our entry point from research exploring the multiple purposes of 'policy pilots' (see Ettelt et al. 2015), recognising the role of pilots in the introduction and testing of new measurements and the OECD's historic use of pilots to overcome opposition to its new assessment frameworks. Ettelt et al. (2015) summarise the purposes of pilots as 'experimentation', 'early implementation' and 'demonstration,' arguing that policy piloting does not primarily serve the purposes of generating evidence through experimentation and evaluation, but rather serves as a policy instrument designed to achieve the objectives of policy-makers rather than to question them. Jowell (2003) likens strategic pilots to kicking issues into the 'long grass' so that the problems will eventually go away and clear the path to wider implementation.

Schleicher (2018) relates how this strategy was employed to overcome resistance to PISA:

> When I proposed a global test that would allow countries to compare the achievements of their school systems… 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. (17-18)

PISA-D was similarly controversial. The Paris Declaration (OECD 2003) saw participants concede that ambitious donor-led interventions had actually drawn down the capacity of aid-recipient nations, vowing that partnership would be country-led and follow domestic priorities. Regarding the SDGs, UNESCO (2014) had initially insisted that "the setting of the

agenda…should not merely reflect indicators that already exist" (4), and that "monitoring and accountability mechanisms should be country-driven" (10). Just five years later, the PISA-D Project Completion Report (OECD 2019a) declared, 'Project implementation was successful.' (OECD 2019b). We suggest that the experimental success of the PISA-D pilot was preordained, while its political success was ensured through attentive stewardship of participating nations throughout the process (see also Addey and Gorur 2020). Consequently, the outcomes of PISA-D replicate rather than address the longstanding critiques of PISA and the extension of ILSAs in low- and middle-income nations.

Our analysis begins with a review of the extant literature which identifies the rationales for extending ILSAs into low-income nations and the longstanding challenges which have limited its realisation; which PISA-D was designed to resolve. We then focus on the PISA-D findings through an analysis of the main OECD report, the six individual country reports, and related seminars and publications. The subsequent section analyses the Educational Prosperity Approach which provided the framework around which PISA-D was developed and interpreted. We then move to examine the extent to which the major benefit of PISA-D, namely learning policy lessons from nations facing similar contexts, was achieved. Prior to concluding we focus on the reduction of grade repetition which was the main policy recommendation that emerged from PISA-D. We argue that PISA-D replicated rather than resolved the challenges it was designed to address and that analysis of its findings, framework and recommendations must extend beyond an appraisal of technical issues to encompass critical reflection on how the pilot initiative serves to advance the OECD's strategic mission (Auld and Morris 2019).

Our analysis connects to existing literature in several respects. We argue firstly, PISA-D reports prioritise the implementation of policies designed to ensure involvement in future PISA cycles. This is central to the OECD's mission, especially by extending the scale

and scope of its assessments (Sellar and Lingard 2014; Lewis 2020). Secondly, rather than using PISA-D to identify 'contextualised' and 'effective policies' (OECD 2018c, 2), policy lessons are largely generic principles framed by the Educational Prosperity Approach (hereafter EPA) and supported by comparisons derived from the main PISA exercise. This finding parallels Lewis' (2017) observations on PISA for Schools, in which he argues the promotion of 'what works' "largely eschews any meaningful consideration of local context, conditions or requirements" (282). Finally, the strategic nature of the pilot aligns with Klees et al.'s (2020) characterisation of the World Bank's science-based interventions as ideology masquerading as evidence, and Schweisfurth and Elliott (2019) more general assertion that global assessment frameworks and policies are "less about transferable certainties and an evidence base than about [global] agencies' desire for influence" (5-6).

We also contribute to the literature on organisational legitimacy by identifying the ways in which the OECD 'gains' authority in monitoring SDG 4, as key agencies and actors compete to extend or, as in the case of UNESCO (Edwards et al. 2018), repair their influence and power (Verger et al. 2018) in the field of global education governance. PISA's legitimacy has largely derived from its portrayal as an objective and universal measurement of educational quality, leveraging the technical expertise necessary to develop and interpret the measurements. The portrayal of new initiatives such as PISA-D, as 'pilot projects', with its connotations of generating evidence through evaluating a policy instrument, similarly legitimates it as an enterprise that has been scientifically proven and can be extended throughout the world.

Extending ILSAs into low-income nations

ILSAs have been positioned as critical for driving educational changes and improving learning in the context of the SDGs (e.g. UNESCO-UIS 2016; World Bank 2018). Although

UNESCO Institute for Statistics (UIS) has stated that there is no single learning metric that can be used for measuring progress to achieve SDG 4, the OECD's strategy on PISA-D promotes PISA as a universal monitoring tool. Tikly (2017) has identified PISA-D as the means of "drawing low-income countries into the same framework of international assessments that is already inhabited by many high- and middle-income countries" (37). This is also observed in Africa by Brock-Utne (2018) where the OECD is seeking to extend its influence. Addey (2017) examines the multiple aims attached to PISA-D, noting that it fits neatly with the Organisation's repositioning as a global organisation and alignment of PISA with the SDGs.

In parallel, the extension of ILSAs into low-income nations has been widely critiqued. Howie (2018), for example, argues that, most of the frameworks underpinning ILSAs to date "have been based upon research in developed countries" and "are insufficient for explaining the variance across and within schools in low- to middle-income countries" (114). This arises because psychological science and development have been based on samples drawn entirely from the WEIRD (i.e. Western, Educated, Industrialised, Rich and Democratic) societies and thus do not represent objective universals (Heinrich et al. 2010). The same issue applies to PISA-D sample questions. According to Kaess (2018), PISA-D "depends on [the main PISA] testing material primarily developed by institutions based in the OECD member states, the majority of which are sovereign states of the Global North" (353).

Scholars have further argued that the issues in low- and middle-income countries are distinctive and need to be dealt with differently (Howie et al. 2017; Wagner et al. 2018). Thus, in 2009, India withdrew from PISA after it performed poorly, and claimed the test was insufficiently adapted to the Indian context. Andreas Schleicher concurred, stating the process was rushed and "implemented without much contextualisation" (Edwards 2019, n.p.),

subsequently crediting Michael Ward with the original idea for PISA-D and outlining the strategic need for pilot nations:

> And we said, "let's try it out in three countries. We need three countries from different parts of the world with good data." And so, I said to him, "Well you know if you want to have three countries, we need to try it out with five. You know, some might not work, and then let's ask seven because some may not agree." So that's what we did… that was important to me, that we don't just test this in one region and one place but that we pretty well cover the world... we had good interest from Latin America, we had good interest from Africa, good interest from Asia." (Schleicher 2019)[3]

The reality appears rather more uneven across participants (see Auld et al. 2019; Addey 2019). Regardless, having secured the required coverage of pilot nations, the OECD commissioned (former) OECD and World Bank employees (Bloem 2013; Lockheed et al. 2015) to review the experience of low- and middle-income countries participating in PISA. The reviews highlighted the significant challenges such nations face, including the lack of resources and institutional capacity, inability to obtain a representative sample of 15-year-olds, and the clustering of students at low proficiency levels.

Gorur et al. (2019) identify the acute challenges which faced those tasked with administering PISA-D and the adjustments needed to deal with local realities, while Addey and Gorur (2020) surveyed the stages through which participants were translated into objects comprehensible through the PISA gaze. Our analysis is informed by these insights and concerns and builds on them by: (i) analysing how 'pilots' operate as a strategic instrument; (ii) presenting and analysing the published findings of the assessment; (iii) examining the underpinning framework (i.e. the EPA) used to collect contextual data and interpret the

---

[3] This version of events contradicts earlier statements by OECD analysts who insisted that the project had been demand-led and initiated by low- and middle-income nations (see Auld et al. 2019).

PISA-D findings, and; (iv) surveying the policy recommendations presented in the main OECD report and six country specific reports.

Findings, cognitive instruments and out-of-school assessments

Detailed accounts of the findings on school-based assessment are provided in the country reports and the OECD main report provides a "comparative overview of the main results" and "where possible, compares these to the OECD average for the PISA 2015 assessment cycle" (2018b, 6). The findings cover: educational attainment at age 15 in reading, mathematics and science; and health, well-being and attitudes towards school and learning. The results are primarily reported through an extended commentary on how the seven PISA-D nations compared with the OECD average in PISA 2015 and with the minimum level of proficiency envisaged in SDG 4 (benchmarked as PISA level 2). The findings are summarised in a table reproduced in Figure 1.

Figure 1

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

PISA-D aimed to bring all nations onto a common scale in line with PISA so that PISA data can be linked more directly to SDG 4.1. This target stresses the minimum proficiency level in reading and mathematics that all children should attain by the end of primary and lower secondary school. The latter was referred to by the OECD as equivalent to Level 2 on the PISA scale. To construct one unbroken developmental trajectory, PISA-D was adjusted within the overall PISA framework and implemented in accordance with PISA's technical standards and practices (OECD 2018a). Panellists at the 2019 International Seminar pronounced adjustment to the PISA-D cognitive instruments on all dimensions a success, though it was confirmed that "there were no new items developed for PISA-D," and that "only existing items were used (OECD 2019b, 4). Instead, PISA-D introduced *more levels* at the bottom end of the PISA scale and created lower proficiency levels (e.g. Level 1a, 1b & 1c). Thus, over 50% of the PISA-D items were used in PISA 2015 and more than two thirds of the items were at Level 2 or below.

Speaking at the Seminar, Ann Kennedy of Educational Testing Service (ETS) reported that the risk that some of the constructs used in PISA cannot be adequately applied, adapted or operationalised in some low- and middle- income contexts to ensure international comparability proved manageable (OECD 2019b, 3). While it is hard to imagine a situation in which a major contractor announced risks could not be managed effectively, Kennedy further affirmed a "strong linkage" between PISA-D to PISA and stated that PISA-D had met the needs of those on the lower ends of the scale. The PISA-D findings thus appear to both address prominent technical concerns and confirm piloting nations' poor performance levels relative to OECD member nations, situating pilot nations on the 'PISA development scale' relative to the basic minimum standards identified by SDG 4.1.1c. Silvia Montoya, Director of the UNESCO UIS, subsequently affirmed the importance of PISA and PISA-D in monitoring global indicator 4.1.1.c (OECD 2019b, 8).

Two key technical problems were pertinent when making adjustments to the PISA scale. First, there are high levels of grade repetition and 'out-of-school' children among the 15-year-olds. In OECD member nations, an average of 89% of 15-year-olds are enrolled in at least grade 7 by age 15, which means they are eligible to sit the PISA tests. In PISA-D pilot countries, however, the OECD main report (2018b) identifies that on average only 43% of all 15-year-olds were enrolled in at least grade 7 by age 15. The remaining 15-year-olds were either in grades below 7 or 'out of school'. This number was lower in Cambodia, Senegal and Zambia, where only around 30% of 15-year-olds were eligible to sit the PISA-D test. Whilst these are presented as technical challenges, their nature and significance suggest that the problems are more fundamental. Even after adjusting the age range from 15 to include 14 to 16-year-olds and those who are 'out-of-school', the OECD still had trouble identifying a suitable population to take the test in out-of-school settings.

The Summary Report from the Seminar (OECD 2019b) notes that to include the out-of-school population in a national assessment results it was necessary to estimate test scores for the whole population (i.e., taking into account those not in school) by putting bounds on unobserved scores. Problems associated with this method of 'imputing' have been raised independently (see Jerrim et al. 2017). The Report indicates Michael Ward described this as "effectively guesswork" and "carried out by assumptions that are not underpinned by real evidence" (5). The outcomes of this experimental dimension of the pilot were carefully qualified. Claudia Tamassia of ETS advised that the adequacy of the PISA frameworks for this unique population should be examined further due to the complexities associated with identifying and surveying participants through existing methods. Still, Ms Leyla Mohaddjer of Westat concluded that the pilot study sampling and survey operations goals were met, and many valuable lessons were learned (6).

Second, the proportion of students who do not speak the language of instruction (and testing) at home varied widely and was far greater than in the standard PISA assessment. The main report notes that "a significant minority of students" in Guatemala (9%) and Paraguay (41%) do not speak the language of instruction at home; and "the vast majority of students" in Senegal (94%) and Zambia (83%) do not speak the language of instruction at home. Not surprisingly, the report notes that pupils "scored significantly higher" in reading when they speak the language of the test at home, "even after accounting for students' socio-economic status and family resources" (OECD 2018c, 9). In common with findings from the regular PISA, this correlated with performances in mathematics and science, with Senegal and Zambia scoring significantly lower than Guatemala and Paraguay. This raises fundamental questions regarding the use of the data to compare education outcomes or maintain alignment with PISA more generally.

When trying to link performance to the various contexts, the main report (OECD 2018c) draws on the questionnaire data which used the EPA as its guiding framework and suggests that caution must be exercised when trying to determine whether certain school or classroom practices have strong relationships with performance. Nevertheless, it declares that "it is possible to identify a range of factors that influence student performance" (OECD 2018c,11), which are initially identified as the EPA's 'prosperity foundations'. In the following section, we provide a closer examination of the EPA, including its origins, how prosperity foundations were identified, the type of developmental model it follows, and its application during PISA-D.

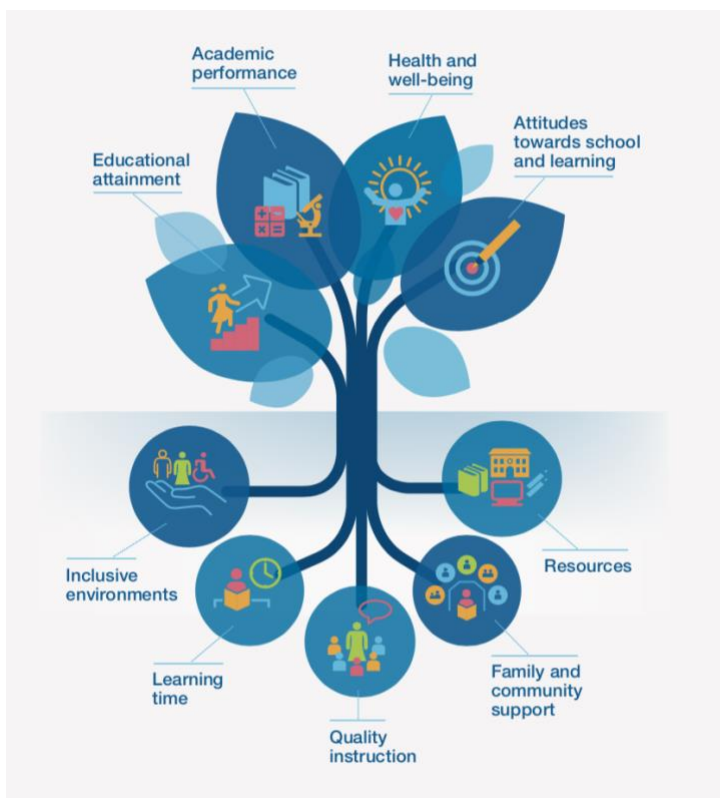**The Educational Prosperity Approach**

Improving the contextual questionnaires was stated as one of the key outputs of the PISA-D pilot (OECD 2019b), which were based on the EPA. The EPA is not a recent development,

but its application to ILSAs such as PISA-D is novel and has largely been overlooked in the literature. The approach was developed by J. Douglas Willms, Founder and President of the Learning Bar – a consultancy that has worked for UNESCO and OECD on aspects of their assessment programmes (Willms 2003; 2006). Willms (2018a) argues that ILSAs have provided evidence on pupils' performance levels, but "in many reports, the analyses based on the 'school effects' paradigm have provided misleading results" (98). These include the OECD reports which have tended to focus on school practices and structures when developing explanations for student outcomes and driving policy lessons (also see Auld and Morris 2016). However, it is virtually impossible to isolate the school effects attributable to many factors that are outside school and operate at a very early age (e.g. biological embedding, parents' engagement, early childhood care etc.). Willms (2018a) thus positions the EPA as: "abandoning the school effects paradigm in favour of collecting rich data on a small set of developmental outcomes and the causal factors that drive these outcomes at several stages of children's development, from conception to adolescence[4]" (99).

The EPA is re-imagined as the 'The Prosperity Tree' shown in OECD publications and is reproduced in Figure 2. Questionnaires for the contextual items in PISA-D were structured according to the EPA's five core 'foundations for success': *inclusive environments*; *quality instruction*; *learning time*; *material resources*; and *family and community support* and these 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*.

---

[4] These include six stages of development: prenatal, early development, pre-primary, early primary, late primary and lower secondary, and upper secondary.

Figure 2 The EPA is re-imagined as the 'The Prosperity Tree'

Whilst the EPA adopts a holistic approach, given its emphasis on the cumulative effects of learning it also relies heavily on assessment results to establish standards and inform progress over an individual's sequential stages of development (i.e. from conception to adolescence). Willms identifies this aspect as essential for supporting the monitoring and achievement of SDG 4:

> The EPA advocates for reliance on monitoring data, as well as a frequent – and early – collection of indicators. These can be used to target policies at the local and national level that bring us closer towards the global education goal (SDG 4). (2018b, para. 2)

The application of the EPA to PISA-D raises several critical issues, notwithstanding Willms' assertion that it had been wholly successful (OECD 2019b). Firstly, as noted earlier, most of the frameworks underpinning ILSAs have been based upon research developed in the global North and are insufficient for explaining variance across low-income nations (Howie

et al. 2017). This equally applies to the EPA, which has its roots in work on pupil literacy in Canada and which assumes that foundations for success are both necessary and universal. As Willms explains:

> The outcomes are considered universal in that they are key markers of child development. Similarly, the Foundations for Success are universal in that a large body of research confirms that they are necessary conditions for success at each stage of development in low-, middle- and high-income countries alike. (2018a, 99)

Secondly, the EPA's prosperity foundations, deemed necessary and universal, assume that higher levels in each 'prosperity foundation' will incrementally enhance 'prosperity outcomes'. Though the relationship is not elaborated, the implication is that those nations that devote more resources to education, provide more inclusive environments, better quality instruction, and greater community and family support, will also enjoy better education outcomes. On the one hand, these are self-evident truisms which derive from the meanings of the words employed. On the other hand, as a developmental model, it replicates *linear* progression models of development; a qualitative variation on the Education Production Functions (EPF's) which dominated education planning during the 1970s and 1980s (Jallade 1970), and which is used by Hanushek (2010) whose analyses of the positive relationship between PISA scores and economic growth underpins the OECD's assessment frameworks.

Finally, the EPA highlights the significance of contextual factors, especially those outside schools, in supporting children's development, but it does not help address the realities or challenges that low-income nations face. As Lockheed (2018) notes, the framework "serves as a valuable reminder of a main purpose of schooling, but it does not address the implementation challenges or resource needs associated with the model" (117). She elaborates:

> In low-income settings – where less than one-fifth of 4-5 year-olds-attend pre-school, first grade sizes of over 100 students are common, internet coverage is low, and many schools lack

electricity – using either of these tools [which involve proprietary software and complex implementation strategies] presents sizeable challenge… Ministers also understand many of the requisites for learning [i.e. the 'foundations for success']. 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. (118)

Despite these concerns, Willms concluded that the contextual questionnaires were delivered successfully to in-school students and out-of-school youth (OECD 2019b). Both the OECD and World Bank have frequently argued that raising pupil performance is primarily a function of how efficiently resources are used in schools, rather than levels of resourcing (e.g. World Bank 2018; OECD 2019c). The PISA-D outcomes underline the importance of resources and access to schooling. Below we analyse the policy recommendations promoted in both the OECD main report and country reports.

Peer Learning and Policy Lessons

PISA-D was presented as an opportunity for peer learning, strengthening political will for reform, and identifying transferrable policies. The concrete recommendations which appear in the main report have two foci. First, PISA-D nations are encouraged to improve their systems for *assessing* pupils as this would allow them to refine and achieve their national education goals and targets:

> PISA-D assessment results provide countries with a solid database that can help them refine policy priorities and set new goals or targets to improve the foundations for success at all levels of their education systems…The challenge for countries over time is to maintain a focus on these goals or targets, and to track progress towards them by participating in future cycles of PISA and other relevant studies. (OECD 2018c, 16)

Second, the most repeated recommendation in the main report is that policies that reduce grade repetition should be prioritised, because it is costly, reduces pupils' attainment and delays their entry into the labour market, as illustrated below:

Grade repetition…is a costly policy, as it requires greater expenditure on education and can delay students' entry into the labour market…Research has found mainly negative effects of grade repetition on academic achievement and attainment. With one-third of students across PISA-D countries reporting that they had repeated a grade, each country should consider replacing grade repetition with practices that have a more positive impact on outcomes. (ibid.)

It is unclear how the PISA-D findings isolate the significance of grade repetition rather than the many other factors (esp. levels of resourcing) which the EPA identifies. The data reported on levels of pupil achievement and the proportion of pupils repeating a grade provides very limited support for that claim: grade repetition in Cambodia and Senegal is significantly higher than in Zambia but pupils' performance in reading, mathematics and science in Zambia is lower than in both those nations. Beneath the EPA's humanitarian framing, the OECD's economic logic is represented in concerns regarding students' timely entry into the labour market.

Similar inconsistencies and contradictions are prevalent both across and within the PISA-D country reports. Table 1 provides a summary of the six country reports[5] and outlines the main problems identified, the recommended key policy actions and the sources of evidence cited to support those policies.

[5] Sources of the original reports are available [here](.).

Table 1 PISA-D National Reports: Summary of problems, recommendations and evidence

Country	Problems identified/prioritized	Key policy recommendations (Short- and medium-term improvements)	Sources of reference/evidence (countries and research/reports cited to support the policies advocated)
Cambodia	**High rate of grade repetition & school dropout**		**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.
		*to **reduce grade repetition***	**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).
			UNESCO (2012): Early support for academically poor students is more cost-effective than giving them a second chance to repeat a grade.
			(OECD, 2016b): Strong support from school improves students' performance in PISA accordingly.
	Low student achievement	*to **ensure quality learning time***	
		(1) strengthen school discipline:	**(OECD, 2013; OECD, 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.
		(2) emphasize classroom assessment	
		(3) increase learning tasks aligned with 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.
		(4) improve the quality of instruction:	**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.**
		(5) improve universal basic skills	**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)**;
Zambia	**Low parental & community involvement**	*to **strengthen family and community support***	
	Lack of resources	*to **improve resource allocation particularly for rural and disadvantaged schools***	**Senegal** has allocated about 23% of total public expenditure towards education.
	Gender & social economic & urban-rural disparities	*to **strengthen teacher recruitment policy***	
		*to **decentralize the textbook procurement***	
	Grade repetition	*to **reduce the school entry age to 6 years:***	All **PISA-D participating countries** have primary school entry age of below 7 years; this was also the trends in the **OECD countries (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)**.
	A learning crisis	*to **strengthen the monitoring and evaluation of formative assessments***	

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 e.g. **Brazil** (SUMMA, 2018a)**, U.S.,** (Bernstein, 2009) and **UK** (Maxwell,2014).
	Gender gap in math performance	to *make changes in classroom and teaching methodology*	**What Works Clearinghouse:** Academic skills of students, esp. girls with lower performance are expandable and improvable. **(Halpern, 2007; Blackwell, 2007):** Teachers are recommended to understand and communicate to their students that math and science skills can be improved through constant effort and learning;
	Inequality & learning differences	to *focus on the distribution and use of educational resources*	**(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 very beneficial, at the level of student performance, and as to compensate for inequalities in education **(Henry et al., 2010)**. **(OECD, 2016a):** in **the highest performing educational systems**, resources are usually distributed more equitably among schools with a high and low socioeconomic index. **The Learning Better (IDB, 2017)** indicates tracking is an effective resource reallocation practice.
		to *focus on teachers and professional development*	**(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. **(What Works Clearinghouse, 2018):** A successful experience that was developed in **the US** to attract high quality teachers is to develop an attractive programme for graduate students with high qualifications in careers not necessarily related to teaching. **(Busso et al., 2017):** Training programmes at work and the use of monetary incentives have proven to be effective, whereas training prior to teaching and selection processes are not.
		to *develop comprehensive evaluation processes and increase the proportion of 15-year-olds who score above PISA level 2*	
Senegal	**Low attendance & quality of education**	to *scale up investment in early childhood education*	
		to *strengthen teacher training through enhanced monitoring mechanism*	
		to *move towards a unified curriculum oriented towards development in PISA skills*	
		to *strengthen the management of the education system*	
	Lack of security in school & noise or disorder in classroom	to *secure school space* **and restore discipline**	
	Low level of educational resources	to *prioritize teachers with modern instructional resources;*	The best performing countries in PISA-D have invested much more in education than others, e.g. **Ecuador** and **Singapore; 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); However, Senegal's effort to reach this amount in the national budget exceeds that of all **PISA-D countries** and even **some OECD countries**;
		to *encourage and pursue decentralization initiatives*	The contribution of parents and parent organisations remains as low as all **PISA-D countries** except **Cambodia.**
Guatemala	**High drop-out rates & grade repetition**	to *establish strategies to support students in transitions*	
		to *review grade repetition policy*	**(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 *provide inclusive environments*	**UN SDG 4 & 7**
		to *increase learning time*	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.
	Lowest performance in Latin America **& low commitment**	to *increase the percentage of students in PISA level 2 or more in reading and science by 10%*	
		to *perform formative evaluation*	**(Hat-tie, 2007):** Formative evaluation, esp. the feedback, is essential to prevent non-learning;
		to *strengthen in-service teacher training and enhance teaching work at the medium level*	
	High report of diseases	to *increase health and well-being*	

Paraguay	**Limited educational resources**	*to increase educational budget and ensure all resources necessary to achieve a quality education*	The accumulated expenditure per student 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.
	low parental involvement	*to increase parental involvement in educational activities and promote informative programmes*	
	Grade repetition	*to double efforts to correct the perception of repetition*	**(Brophy, 2006):** Evidence accounts for the harmful effects of repetition on learning and on the efficiency of public spending on education.
		to review the repetition education policy	
		to implement remedial strategies for students of basic and secondary education	
		to enhance early childhood education:	*(Willms, 2015):* children who live in a context that contributes to their training, with sufficient resources to learn, receive stimuli that are up to their abilities and are consistently accompanied by an adult during their growth, reach higher levels of development. *(UNICEF, 2017):* 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.
		to design and implement a structured and well-detailed curricular programme	
	Low quality of instruction	*to install the culture of formative and systematic evaluation*	**(MEC, 2013):** «Teaching Sequences» offers standardized instruments that allow the identification of each student's progress according to performance levels and areas of competence.
		to improve teacher training system	**(Burns and Luque, 2015):** improved teacher training systems, together with the improvement of work incentives would contribute to the hierarch of the teaching career, attracting the best students and professionals.
	Low quality of learning	*to make effective use of time*	**(Alfaro, Evans and Holland, 2015):** 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.
	Intensity of violence	*to mitigate the intensity of violence*	
	Insufficient family & community support	*to strengthen family and community support towards school activities*	**(Avvisati et al., 2014; Berlinski et al., 2016; Cerdan-Infantes and Filmer, 2015; Bauer et al., 2018):** 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. **(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.

At the outset, there is a notable divergence of policy recommendations across the countries, which suggests different approaches to and some domestic influences on the construction of the national reports. While the reports are the responsibility of the relevant national Ministries, they were strongly influenced by the OECD; the process is explained thus:

> With the support of the OECD, each country has formed an analysis task force and national working group, and has nominated one analyst to lead the work on data analysis and reporting. The OECD works directly with the analysis task force, especially the lead analyst from each country. Indeed, the OECD implemented a lead analysts' programme from September 2017 through December 2018. Its aim was to help lead analysts complete their analysis of PISA-D data and prepare a national report. (OECD 2018b, 1-2)

The result is that, some, such as the Latin American countries, refer to the EPA but are largely framed with regard to their own national plans and priorities. In contrast, the Cambodia and Zambia reports adhere tightly to the OECD narrative and mirror the main report's focus on grade repetition, student drop-out and the effect of grade repetition in delaying entry to the labour market.

Here we highlight three critical aspects of the country specific policy recommendations, which relate to: (i) the policies promoted, (ii) the evidence cited, and (iii) the role of the EPA.

Firstly, among the wide array of policies promoted to improve learning outcomes, the main report's emphasis on the need to enhance assessment capacity and reduce grade repetition is echoed in most of the country reports. For example, regarding assessment capacity, the Ecuador report states:

> In order to understand if we are moving towards these objectives and to what extent we are doing so, we need comprehensive, reliable and rigorous evaluation processes …Constant evaluation and participation in international programmes will help us to closely follow our trajectory towards the fulfilment of the goals of the National Development Plan and the SDGs. (2018, 135)

The Guatemala report goes further, asserting that: "one of the practices of the countries with the most effective educational systems is their periodic participation in international evaluations" (2018, 213). Regarding grade repetition, the Cambodia report positions this as 'a low-hanging fruit' which can "bring about substantial changes in student learning at the least expense of finance" (2018, 135). In contrast, the focus on reducing grade repetition is absent in the report of Senegal, where grade repetition is actually the highest (50%) among PISA-D countries. In Zambia, the benefits and barriers to reducing grade inflation are noted and it suggests that,

> The Directorate of Planning and Information may rethink the school entry policy by ensuring that: (i) all children who attain the school entry age are enrolled in school and (ii) reducing the school entry age to six years as learners could have already been exposed to schooling before the age of six through early childhood education. (2018, 127)

Secondly, the evidence cited to support policy recommendations draws on (i) previously published materials and (ii) references to practices which are deemed to work elsewhere. There is a marked diversity across the reports on both these dimensions. Cambodia is at one extreme: the references to published evidence rely heavily on OECD reports and references to nations which perform well on the main PISA test**.** The Cambodia report asserts that there is a need to align learning tasks with the PISA test items and that this is a feature of successful Asian countries (South Korea, Hong Kong, Macao, Vietnam and Taiwan). In contrast, the Ecuador and Senegal reports draw on a wider range of published sources and reference countries. Paraguay was distinctive in that it made no references to OECD reports and cited both a diverse range of research articles, many from regional sources, and evidence from national assessments. Only Senegal made any reference to other nations participating in PISA-D and those references were solely used to argue for increased expenditure on education.

The prevalence of references to policies in high-performing PISA nations suggests that, while PISA-D "allows participating countries to determine whether their policies differ *from those of countries with a similar social and economic context"* (2018b, 11, our emphasis added), policy lessons have been drawn from OECD countries and high-performing systems. Given the reliance on high performers on PISA and the nations in the Global North as the source of evidence, the policy recommendations are very loosely linked to the data and could have been advocated without undertaking PISA-D. One glaring absence from both the main report and the country reports is the failure to discuss the policy implications of arguably the major source of differential performance; namely, the proportion of students for whom the language of the test was not their mother tongue. However, unlike grade repetition, that does not have a direct effect on pupils' eligibility to participate in PISA, or delay entry to the labour market.

Thirdly, the reports adopt the thematic strands and terminology of the EPA to identify their problems and the corresponding solutions:

> Each report, containing six chapters, is based on a template prepared by the OECD and includes analyses of the four key outcomes, known as "Prosperity Outcomes", for each stage of schooling and child development, and the "Foundations for Success" that underpin them. (OECD 2018b, 1-2)

As noted earlier, overall the policy recommendations framed by the EPA lack specificity and focus on factors which relate most directly to schooling, reflecting the OECD's tendency to focus on educational systems and schools as the locus for improving learning outcomes rather than on the broader factors and cumulative effects inherent in the EPA. For example, in Cambodia, the report recommends that low student achievement should be addressed by "ensuring quality learning time" (2018, 137). The Guatemala report emphasises the need to "provide inclusive learning environments" and "increase learning time" (2018, 214-234). And in Paraguay, the report emphasises the need to enhance "quality

of instruction" and "make effective use of learning time" (2019, 186-189). The few references to the EPA's concern for factors outside schools appear in the recommendations to: "strengthen family and community support" (Zambia and Paraguay) and 'parental involvement' (Paraguay).

In this way, the entire undertaking, from the collection of data to its translation into reform priorities, is largely framed by a prefabricated developmental model. References appear haphazardly both across and within the individual reports and it is readily apparent that the framework has not been applied consistently or with shared understanding. Specific interventions are then selected from a range of sources, but not from comparison with policies and practices in other PISA-D pilot countries as the OECD had indicated. Although the problems they face are markedly different, policy lessons for PISA-D nations were derived from more economically developed nations. Overall, the nature of the recommendations reflects the OECD's drift towards 'industrial benchmarking' (OECD 2014b) and 'data-inspired speculation' (Valverde 2014), whereby its assessments are positioned as the central barometer of quality, but policy lessons eschew claims on causality and are largely limited to abstracted lists of what successful and improving nations are doing.

Revisiting Grade Repetition

Before concluding, we revisit the central policy recommendation repeatedly advocated across the reports: reducing grade repetition In essence it is advocated that if more 15-year-olds are enrolled at the appropriate grade level this is a prerequisite condition for moving on to the next stage of the EPA– achieving at least minimum levels of proficiency in key subjects. In practice this requires ensuring that pupils are eligible for sitting the PISA/PISA-D test.

The main report's assertion that "research has found mainly negative effects of grade repetition on academic achievement and attainment" (OECD 2018c,16) may be accurate, but

the evidence for supporting that is primarily derived from OECD studies which have focused on the performance of pupils who repeat a grade (e.g. OECD 2011). The situation in India, where policy decisions relating to grade repetition were strongly influenced by claims about its impact on the performance of all pupils, provides an opportunity to examine the complexities of such policy claims.

In 2009, the 'Right to Education' Law mandated that no child can be detained or held back in a class until the completion of elementary education up to class grade 8. The provision, referred to as the 'No detention Policy' (NDP), was introduced to retain those children who used to drop out and promote a joyful and fear free school environment by respecting children's pace of learning. Following concerns as to the impact of NDP, especially on educational standards and levels of pupil and teacher motivation (Sabharwal 2018), two Government Committees were convened from 2015 which were critical of the policy. Subsequently, the 28 States were consulted and 23 requested an amendment of NDP. In 2017, the central government proposed a bill which would allow pupils to be examined at grades 5 and 8 and held back if they failed a re-examination. In 2018, the Law was passed, and it was left to the 28 States to decide how to implement the new policy, which effectively scrapped the NDP. This has been highly contested, for example, Taneja (2018) argues that the evidence does not support the abolition of NDP:

> Poor quality of education and declining learning outcomes are inevitable in a school system where 50% schools lack headteachers, 8% of primary schools have only one teacher and 90% schools lack the minimum infrastructure laid down by the law. Good teaching requires teachers to be qualified, trained, motivated and supported. (3)

The point of this excursion is not to argue for or against grade repetition but rather to highlight the problem of providing a selective portrayal of a specific policy and its relation to education outcomes. As the above quotation highlights, if the benefits of reducing grade repetition are to be realised, this requires substantial investment of resources, especially in

teachers, curricula and remedial support for those pupils affected. Such interventions are not identified as a priority across the various PISA-D reports. The value of reducing grade repetition, and involvement in international assessments more generally, is also not widely promoted by other studies that compare the economic returns of policies promoted in low-income nations. For example, the Copenhagen Consensus Centre (2014) claim that investing in family planning generates the highest returns and they only identify one educational policy, the provision of pre-school education, among those which generate a high ratio of benefits to costs. Those do not feature in PISA-D.

Conclusion: from proof of concept, to proof of progress

PISA-D aimed to overcome well-established barriers preventing low- and middle-income countries from engaging in PISA. By introducing new proficiency levels at the lower end of the scale and anchoring PISA to the basic minimum standards identified under the UN's education SDG 4.1, the OECD has indeed been able to position nations at the lower end of the PISA scale. The pilot did serve a limited experimental purpose, with technical adjustments made to both cognitive instruments and questionnaires. At the same time, these were contained within the wider OECD assessment frameworks. Whilst technical experts sincerely engaged with the issues, the success of this aspect of the PISA-D pilot was preordained as the pilot was deployed as a policy instrument to help achieve the OECD's objective of extending PISA 'throughout the world'. With the proof of concept validated, PISA would provide proof of progress for the international community.

The results were performed through national reports and public seminars, bringing together OECD officials and mainly those engaged in the project who vouched for its success. Clearly this is not empty theatre. Noting that pilot nations have largely signed up to participate in PISA by 2021 or 2024, and stressing the political nature of the exercise, Addey

and Gorur (2020) "deduce that the OECD and the nations participating in the PISA-D pilot believe that the PISA-D pilot was successful in achieving its aims" (2). Indeed, there are many reasons political elites of a given nation decide to participate in ILSAs (Addey and Sellar 2018). The pilot process is consistent with Grek's (2013) account of the OECD's policy work in Europe, whereby the Organisation carefully manages technical experts and political elites to forge consensus, nurturing a transnational epistemic community which Grek refers to as the 'expertocracy'. Pettersson and Popkewitz (2019) note that 'expertism' is something that has constantly to be seized and debated, observing one of the main impacts of PISA has been to place the expertise in education in the hands of entrepreneurs, technicians, economists and statisticians.

Once a nation signs up and leaders are commended for their 'courage' (Schleicher 2019), opting out becomes politically problematic, particularly once the assessment is anchored to a global agenda and tied to the disbursement of aid (OECD 2013, 2019a). This is significant given that participation in PISA will redefine perceptions of what is important in education, setting reform priorities and guiding the allocation of scarce resources. Overall, PISA-D extends what Nóvoa (2018) terms the 'solutionist drift' in comparative education, which he describes as: "…an uncritical appropriation and generalisation of global solutions imposed by data and evidence on 'what works'…based on the false idea of consensus on the aims of education and the means for achieving them" (551). He notes, in terms which echo the depiction of PISA as a form of 'banal imperialism' (Silova and Auld 2019), that both the solutions and means are "always found in the examples, images, and models of the Anglo-Saxon North, never on Planet South" (553). Still, from the national reports it is apparent that the assessment will be mobilised towards diverse ends.

This opens into more fundamental issues regarding the value of engaging in PISA, which have been well-rehearsed in the literature (e.g. Cowen et al. 2011; Meyer and Benavot

2013) and PISA-D embodies the same established problems inherent in PISA. Andreas Schleicher recently acknowledged three major limitations with the assessments. First, he questioned whether PISA and PISA-D is still relevant:

> 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 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 cont*exts. 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. (Li and Auld 2020, 11)

Second, he questions PISA's capacity to identify policy lessons. In *World Class*, Schleicher (2018) states 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" (61). As he notes, "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.). Finally, he outlines changes to PISA that undermine its positioning as a barometer of progress. Specifically, Schleicher (2018) notes the OECD's decision for the assessment to "constantly evolve" rather than function as a "fixed point", a decision that was taken to allow PISA to "lead education reform" by "measuring students what they will need to thrive in the future" rather than "what was considered important some point in the past" (277). This rudderless steering seems like considerable power given the Organisation's democratic deficit, though it does enable the OECD to dismiss critique as time-lapsed while adding fresh layers of data and reinventing assessments based on changing visions of the future. In short, evolution and uncertainty are an effective strategy for safeguarding legitimacy and therefore organisational survival.

This conclusion is not novel. Nonetheless, it remains important. International organisations compete and cooperate to forge world order (Elfert 2017), a process pursued in part by aligning strategic narratives with competing organisations and partners. The legitimacy of these narratives increasingly relies on the construction of calculable worlds that render abstract ideals visible and amenable to control. The OECD leverages its technical expertise as statistical legitimacy and contributes to the symbolic organisation of the world into a coherent form of order. In this respect, it is successful. At the same time, stories that do not fall within the assessment's mental map of the world are ignored. That is, it matters what kinds of stories we tell, what thinking our data makes possible and what futures are precluded. Notably, whilst the OECD's visions of the future stress sustainability, the rationale for improving PISA scores remains to maximise economic growth while the finiteness of the world's resources is ignored (Rappleye and Komatsu 2020). We suggest the OECD might direct some of its technical expertise to pilots that provide statistical substance to other stories, adding to rather than subtracting from the manifold realities of the global and opening alternative possibilities for the future. PISA-D failed in this most urgent task, but its success suggests it is just a symptom of a much wider collapse of the collective imagination.

References

Addey, Camilla. 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 editor

Antoni Verger, editor Mario Novelli, and editor Hülya K. Altinyelken, Second edition. London: Bloomsbury Academic.

Auld, Euan, and Paul Morris. 2016. 'PISA, Policy and Persuasion: Translating Complex Conditions into Education "Best Practice"'. *Comparative Education* 52 (2): 202–29. https://doi.org/10.1080/03050068.2016.1143278.

———. 2019. 'Science by Streetlight and the OECD's Measure of Global Competence: A New Yardstick for Internationalisation?' *Policy Futures in Education* 17 (6): 677–698. 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.

Benavot, Aaron, and Heniz-Dieter Meyer. 2013. *PISA, Power, and Policy: The Emergence of Global Educational Governance*. https://doi.org/10.15730/books.85.

Bloem, Simone. 2013. *PISA in Low and Middle-Income Countries*. OECD Education Working Papers; No. 93. Paris: OECD Publishing. http://dx.doi.org/10.1787/5k41tm2gx2vd-en.

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.

Copenhagen Consensus Center. 2014. 'Benefits and Costs of the Population and Demography Targets for the Post-2015 Development Agenda'. Copenhagen Consensus Centre. https://www.copenhagenconsensus.com/publication/post-2015-consensus-population-and-demography-assessment-kohler-behrman.

Cowen, Robert, Hans-Georg Kotthoff, Miguel A. Pereyra, and Comparative Education Society in Europe. 2011. *PISA under Examination: Changing Knowledge, Changing Tests, and Changing Schools*. Comparative and International Education. Rotterdam: Sense. http://dx.doi.org/10.1007/978-94-6091-740-0.

Edwards, D. Brent, Taeko Okitsu, Romina da Costa, and Yuto Kitamura. 2017. '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*, 2019. https://www.devex.com/news/sponsored/india-s-re-entry-to-pisa-triggers-mixed-response-94286.

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.

Elfert, Maren. 2017. 'UNESCO, the World Bank, and the Struggle over Education for Development through the Lens of the Faure Report (1972) and the Delors Report (1996)'. *L'Education En Débats: Analyse Comparée*, 5–21.

Gorur, Radhika, Estrid Sørensen, and Bryan Maddox. 2019. 'Standardizing the Context and Contextualizing 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. 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.

Hanushek, Eric A. 2010. 'The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes'. OECD. https://doi.org/10.1787/9789264077485-en.

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.

Howie, J. Sarah. 2018. '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 D. A. Wagner, S. Wolf, and R. F. Boruch. UNESCO/ IIEP.

Howie, Sarah J., Vanessa Scherman, and Surette van Staden. 2017. 'Frameworks for the Analysis of Monitoring Data Related to Achievement'. *Monitoring the Quality of Education in Schools*, 159–94.

Jallade, Jean-Pierre. 1970. 'Educational Planning Methods. Background Study No. 8. Conference on Policies for Educational Growth. (Paris, France, June 3-5, 1970.)'. *OECD Publishing*. https://eric.ed.gov/?id=ED069023.

Jerrim, John, Luis Alejandro Lopez-Agudo, Oscar D. Marcenaro-Gutierrez, and Nikki Shure. 2017. 'What Happens When Econometrics and Psychometrics Collide? An Example Using the PISA Data'. *Economics of Education Review* 61: 51–58. https://doi.org/10.1016/j.econedurev.2017.09.007.

Jowell, Roger. 2003. 'Trying It Out: The Role of "Pilots" in Policy-Making: Report of A Review of Government Pilots'. London: Cabinet Office, Strategy Unit.

Kaess, Kathleen. 2018. 'The Cross-Atlantic Knowledge Divide, or PISA for Development: Should One Size Ever Fit All?' *Atlantic Studies* 15 (3): 349–364. https://doi.org/10.1080/14788810.2017.1370356.

Klees, Steven J., Mark Ginsburg, Hiba Anwar, Melanie Baker 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.

Lewis, Steven. 2017. 'Governing Schooling through "What Works": The OECD's PISA for Schools'. *Journal of Education Policy* 32 (3): 281–302. https://doi.org/10.1080/02680939.2016.1252855.

Lewis, Steven. 2020. 'Providing a Platform for "What Works": Platform-Based Governance and the Reshaping of Teacher Learning through the OECD's PISA4U'. *Comparative Education* 0 (0): 1–19. https://doi.org/10.1080/03050068.2020.1769926.

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* 0 (0): 1–19. https://doi.org/10.1080/03050068.2020.1781397.

Lockheed, Marlaine. 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 D. A. Wagner, S. Wolf, and R. F. Boruch. UNESCO/ IIEP.

Lockheed, Marlaine, Tijana Prokic-Bruer, and Anna Shadrova. 2015. 'The Experience of Middle-Income Countries Participating in PISA 2000-2015'. Paris: Organization for Economic Co-operation and Development. http://www.oecd-ilibrary.org/content/book/9789264246195-en.

Nóvoa, António. 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.

OECD. 2003. 'Rome Declaration on Harmonisation'. OECD Publishing. https://www.oecd.org/dac/effectiveness/31451637.pdf

———. 2011. 'When Students Repeat Grades or Are Transferred out of School: What Does It Mean for Education Systems?' PISA IN FOCUS. Paris: OECD Publishing. https://www.oecd.org/pisa/pisaproducts/pisainfocus/48363440.pdf.

———. 2013. 'PISA for Development General Presentation'. Presented at the Initial Technical Meeting, Paris. http://www.oecd.org/pisa/aboutpisa/initialtechnicalmeeting.htm.

———. 2014a. 'Call for Tender 100000990 - PISA for Development Strand A and Strand B'. Paris. http://tinyurl.com/prjry24.

———. 2014b. 'Measuring the Digital Economy: A New Perspective'. OECD. https://doi.org/10.1787/9789264221796-en.

———. 2018a. *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.

———. 2018b. 'PISA for Development Brief 25: Strengthening Analysis and Reporting in PISA for Development Countries'. OECD Publishing. https://www.oecd.org/pisa/pisa-for-development/25_Strengthening_analysis_reporting_in%20PISA_D_countries.pdf.

———. 2018c. 'PISA for Development: Results in Focus'. https://www.oecd-ilibrary.org/education/pisa-for-development_c094b186-en.

———. 2019a. 'PISA for Development Project Completion Report'. http://www.oecd.org/pisa/pisa-for-development/PISA%20for%20Development%20Project%20Completion%20Report.pdf.

———. 2019b. 'Summary Report of the PISA for Development International Seminar'. http://www.oecd.org/pisa/pisa-for-development/Summary_record_PISA-D_International_Seminar.pdf.

———. 2019c. 'Working and Learning Together: Rethinking Human Resource Policies for Schools'. OECD Reviews of School Resources. Paris: OECD Publishing. https://doi.org/10.1787/b7aaf050-en.

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 editor Christina Mølstad and editor Daniel Pettersson, 18–36. Abingdon, Oxon: Routledge.

Rappleye, Jeremy, and Hikaru Komatsu. 2020. 'Towards (Comparative) Educational Research for a Finite Future'. *Comparative Education* 56 (2): 190–217. https://doi.org/10.1080/03050068.2020.1741197.

Sabharwal, Karan. 2018. 'No Detention Policy: Rethinking Education System of India'. *International Journal of Academic Research and Development* 3 (1): 609–14.

Schleicher, Andreas. 2018. 'World Class: How to Build a 21st-Century School System'. OECD Publishing. https://www.oecd-ilibrary.org/education/world-class_9789264300002-en.

———. 2019. 'Principal Drivers and Rationale for Enhancing PISA in the Context of SDGs'. Presented at the PISA for Development International Seminar, London. http://www.oecd.org/pisa/pisa-for-development/pisafordevelopmentinternationalseminar.htm#:~:text=The%20PISA%20for%20Development%20International,discuss%20the%20outcomes%20and%20results.

Schweisfurth, Michele, and Julian Elliott. 2019. 'When "Best Practice" Meets the Pedagogical Nexus: Recontextualization, Reframing and Resilience'. *Comparative Education* 55 (1): 1–8. https://doi.org/10.1080/03050068.2018.1544801.

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): 917–936. https://doi.org/10.1002/berj.3120.

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*, no. 2 (October): 182.

Taneja, Anjela. 2018. 'No Detention under the RTE Act: The Policy Options'. *Education Policy Brief,* The Policy Options (1). https://www.academia.edu/37058231/No_Detention_under_the_RTE_Act_The_Policy_Options.

Tikly, Leon. 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.

UNESCO. 2014. 'Position Paper on Education Post-2015'. Paris. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/ED_new/UNESCOPositionPaperOnEDpost-2015.pdf.

UNESCO Institute for Statistics. 2016. 'The Cost of Not Assessing Learning Outcomes'. http://uis.unesco.org/sites/default/files/documents/the-cost-of-not-assessing-learning-outcomes-2016-en_0.pdf.

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.

Wagner, D. A., S. Wolf, and R. F. Boruch, eds. 2018. *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries*. UNESCO/ IIEP.

Willms, J. Douglas. 2003. 'Student Engagement at School: A Sense of Belonging and Participation - Results from PISA 2000'. Paris: OECD Publishing. http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33689437.pdf.

———. 2006. *Learning Divides: Ten Policy Questions about the Performance and Equity of Schools and Schooling Systems*. UNESCO Institute for Statistics. https://unesdoc.unesco.org/ark:/48223/pf0000147066.

———. 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 D. A. Wagner, S. Wolf, and R. F. Boruch. UNESCO/ IIEP.

———. 2018b. 'Educational Prosperity: Looking Beyond Equality to Equity'. UNESCO Institute for Statistics. *Data for Sustainable Development* (blog). 2018. https://sdg.uis.unesco.org/2018/10/09/educational-prosperity-looking-beyond-equality-to-equity/.

Verger Antoni, Novelli Mario, and Altinyelken K. Hülya. 2018. *Global Education Policy and International Development: New Agendas, Issues and Policies*. Second edition. London: Bloomsbury Academic.

World Bank. 2018. 'World Development Report 2018: Learning to Realize Education's Promise'. https://www.worldbank.org/en/publication/wdr2018.