

Seeing like PISA: A cautionary tale about the performativity of international assessments

Radhika Gorur

Deakin University, Australia

Abstract

PISA is an extremely influential large-scale assessment, and its ‘policy lessons’ are being incorporated in a range of nations all over the world. In this paper I argue that not only is PISA influencing policies and practices, but also that ‘seeing like PISA’ is becoming a widespread phenomenon. Globally, education administration is now characterized by an intense focus on output measurement, a highly competitive environment heightened by national and international rankings, and an economic and instrumentalist approach to education and education reform. Using James Scott’s account of 18th Century German forestry practices as a parable, this paper suggests that ‘seeing like PISA’ could have far reaching and damaging effects. The paper proposes the following: first, understanding PISA as a ‘project of legibility’ enhances our appreciation of its purposes and possibilities. Second, PISA is much more than a ‘representation’ of existing conditions, but is creating new conditions – in other words, it is not descriptive but performative; and, finally, ‘seeing like PISA’ is bringing about deep-rooted changes, and it is likely that the effects will be very long-term. Some of these effects may only manifest themselves in the next fifteen or twenty years; and, by then, the possibilities of redressing some of the ill effects may be very limited.

Keywords

Education policy, international assessments, performativity, PISA, STS

If you Google® ‘PISA’, the first entry that shows up may not be the Italian city which has been in existence since the 5th Century BC but, quite possibly, the OECD’s education survey, the Programme for International Student Assessment (PISA), which was started in the late 1990s. The first PISA survey was conducted as recently as 2000. In the 16 years since then, participation has

Corresponding author:

Radhika Gorur, Deakin University, 221 Burwood Highway, Melbourne, Victoria, 3125, Australia.

Email: radhika.gorur@deakin.edu.au

more than doubled; PISA participants now represent 90% of the world's Gross Domestic Product (GDP). A report on the policy impact of PISA found 'over 85 per cent of policy makers, local government officials, academics and researchers report having a relatively high level of knowledge of PISA processes and impact' (Hopkins et al., 2008: 19). PISA rankings, released simultaneously around the world every three years, make media headlines globally. PISA is the largest component of the OECD's education budget. Countries like Australia have introduced the ambition of being in the 'top five' in PISA rankings (Gorur and Wu, 2015). Breakspear (2012: 4) found that 'PISA results have had an influence on policy reform in the majority of the participating countries/economies'.

However, this paper is not only concerned with the extraordinary and growing influence of PISA. It is also about the phenomenon of 'seeing like PISA' – that is, the gaining of a reductionist, synoptic and largely economic view that has been facilitated by the development of international indicators and large-scale comparative assessments. Furthermore, it is about the possibilities and dangers of creating a world in the image of the reductionist view afforded by such measurements (Knorr-Cetina, 1999; Woolgar, 1991). Increasingly, I argue, countries are beginning to 'see like PISA'.

To be clear, I am not claiming that 'seeing like PISA' is caused solely by PISA. PISA itself would not have been viable without the infrastructure that took decades to develop: the international indicators, the statistical institutes, the early work in global literacy assessments, the advances in psychometrics and the development of such techniques as Item Response Theory, to name just a few of the key actors in this assemblage (Gorur, 2011; Gorur, 2014). Nor would PISA have been able to gain such popularity without a widespread social policy environment characterised by standardisation, competition, choice, audit and accountability. PISA is at once a product of such an assemblage, as well as a key actor in its perpetuation.

To elaborate the phenomenon of 'seeing like PISA' and to reflect on its consequences, I use James Scott's account of 18th Century German forestry management techniques, and their extremely influential and ultimately disastrous attempts to domesticate and regulate nature, as a cautionary tale. In his seminal book *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* (Scott, 1998), Scott chronicles the rapid growth of scientific forestry practices developed between 1765 and 1800 in Prussia and Saxony, and the destruction caused by these practices in the long run. Using this as a parable, Scott describes several grand schemes that were developed to tame societies and nature to produce social and natural order, which also ended in disaster. Scott's concern was not that the schemes to make things orderly and governable had failed; rather, the schemes had succeeded all too well in this aspect of their endeavour. Where they had failed was in improving the human condition, bringing about, instead, hardship and possibly irreversible damage. In this paper, I use this story as a parable to explore the current schemes to develop global education metrics and render education globally comparable on international league tables, using the example of PISA.

The use of parable in academic writing is allied to the same logic as the use of example (Massumi, 2002). Massumi sees the use of example as experimentation – one that activates detail – and it is through the details that we might encounter and connect unexpected concepts. Serres, making reference to the 'constant use of parables in the Gospel', asserts that, 'Philosophy can be summed up in little stories' (Serres and Latour, 1990). The use of parable or story, and indeed other similar genres, such as Latour's (1996) 'whodunit' ARAMIS, or Law's (2002) performance of fractional coherence in *Aircraft Stories*, is more than a matter of pushing the boundaries of academic writing styles or avoiding jargon to communicate an idea economically or to a diverse audience. Unlike more conventional academic writing which pushes us towards singularities and requires us to fix, conclude and close out possibilities, the use of example, allegory and parable offers a means of generating new and unexpected ways to think about phenomena, and 'a way of knowing the multiple and the

ambivalent' (Law, 2004: 153). Using allegory, parable and other generative genres, writing becomes not merely a communication of facts and findings, but a performance that seeks to make a difference. In explicitly setting this paper up as a parable, or a 'cautionary tale', I recognise the acts of writing and researching as performances of ontological politics (Mol, 1999).

This paper proposes the following. First, understanding PISA as a 'project of legibility' and as a way of seeing enhances our appreciation of the pervasive possibilities afforded by such projects. Second, PISA is much more than a way of representing existing conditions. It is creating new conditions – in other words, it is not just descriptive but also performative. And, finally, the changes 'seeing like PISA' is bringing about are deep-rooted, and it is likely that the effects will be very long-term. Some of these effects may only manifest themselves in the next 15 or 20 years, and by then, the possibilities of redressing some of the ill effects may be very limited.

German forestry and PISA: an uncanny parallel

Scott uses the story of the German forestry management techniques developed between 1765 and 1800 as a parable to explore large-scale projects of imposing order on complex social worlds to make them more politically and administratively convenient to govern. Taking an exclusively fiscal view, forestry management practices in Prussia and Saxony focused single-mindedly on increasing the timber yield. A series of measures was put in place so that the forest became more legible and more readily monitored and more easily administered, and the yield more reliably predictable. Eventually, the abstract forests of the fiscal ledger books came to be replicated in real-life, through such measures as single-species and simultaneous planting in geometrically precise rows. Timber yield increased and it became easier to predict the yield. The spectacularly successful German scientific forestry practices were taught in universities and widely adopted around the world. However, the singularly fiscal understanding of the forest to the exclusion of all the other aspects of forest ecology proved to be disastrous for forestry in the long run. The practice of monoculture depleted the soil capital and the bio-diversity of the forests and made the forests susceptible to a variety of threats; but it was only about 80 years later that the adverse effects of these practices began to be apparent.

There is an uncanny parallel between the vision and practices of Prussian forestry management and the effects they had on forestry practices globally, and the phenomenon which I am describing as 'seeing like PISA'. The sections that follow elaborate these parallels.

A fiscal view

Forests in Prussia and Saxony were assets that were highly valued by the Crown. Timber was required for shipbuilding and construction, industries crucial for a state's economic security. As a result, measuring the amount of timber, predicting the yield and managing the resource to ensure high yields and a steady supply was critical to the state. Forests were therefore almost exclusively viewed through a fiscal lens, and the interest was in the single number represented by the annual timber yield. 'Nature' was translated into a 'natural resource'.

Education is now similarly being regarded almost exclusively in economic terms, as a means of increasing a nation's 'human capital' and raising the GDP. Measuring the education system's 'outcome' and predicting the economic 'yield' from school systems is now seen as crucial to states' fiscal interests. Some of the most influential data on education systems globally are being generated by an economic organisation – the Organisation for Economic Cooperation and Development (OECD). Among OECD's key missions, as described on its website, are to help 'governments around the world' to '[r]estore confidence in markets and the institutions that make them function'

and to '[e]nsure that people of all ages can develop the skills to work productively and satisfyingly in the jobs of tomorrow'.¹

So close has this association between education and the economy become that PISA points are now equated with Gross Domestic Product (GDP). If in German forestry nature had been translated into a 'natural resource' through valuation and measurement, PISA and other such measurements provide the calculations to translate humans into 'human capital'. Expressed in numbers, human capital is easily linked to the demands of labour, productivity and ultimately the nation's future wealth. This excerpt from the OECD volume *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education* illustrates this widely used economic rationale for improving education performance:

In this globalised world, people compete for jobs not just locally but internationally. The integrated worldwide labour market means that highly-paid workers in wealthier countries are competing directly with people with much the same skills but who demand less compensation in lower-wage countries. The same is true for people with low skills. The competition among countries now revolves around human capital and the comparative advantage in knowledge. (OECD, 2011: 14)

PISA presents trend data, analysing gains and losses in performance over time. Such calculations also provide ways to predict the yield from education in terms of GDP, in much the same way as the timber yield from forests was predicted in German forestry management. A study by Hanushek and Woessmann, published by the OECD, calculates the effects on the GDP of raising PISA scores:

A modest goal of having all OECD countries boost their average PISA scores by 25 points over the next 20 years ... implies an aggregate gain of OECD GDP of USD 115 trillion over the lifetime of the generation born in 2010 (as evaluated at the start of reform in terms of real present value of future improvements in GDP).... Bringing all countries up to the average performance of Finland, OECD's best performing education system in PISA, would result in gains in the order of USD 260 trillion.... (OECD, 2010: 6)

What is particularly interesting here is that a causal connection is assumed between increasing PISA scores and increasing GDP. As a result, increasing PISA scores comes to be viewed as a strategy for increasing GDP. However, there is no proof as yet with regard to whether students with higher PISA scores will go on to contribute to raising nations' GDPs, even if one were to accept the fiscal logic of this project. The calculations in the OECD report are based on correlations between a nation's GDP *at the time* and its PISA performance. The first set of students who undertook PISA in 2000 are, in 2016, now barely 30 years old. At this stage, they would have had few opportunities to affect significantly a nation's GDP. Whether or not PISA scores make a difference remains to be seen.

Narrowing the field of vision

Although forests were used in multiple ways by those living near them, it was the timber in the forests, the most fiscally relevant part of the forest, that became the focal point of the state's calculations. Like the proverbial inability to see the woods for the trees, Scott explains, the utilitarian state could barely see the 'real, existing forest for the (commercial) trees'. Not only were the vast majority of the flora and fauna absent in the state's calculations, the radically narrowed vision also 'typically ignored the vast, complex, and negotiated social uses of the forest for hunting and gathering, pasturage, fishing, charcoal making, trapping, and collecting food and valuable minerals as well as the forest's significance for magic, worship, refuge, and so on' (Scott, 1998: 13). This reduction of the forest into a narrow focus on timber yield left out an extremely wide range of things that were integral to the forest:

Missing, of course, were all those trees, bushes, and plants holding little or no potential for state revenue. Missing as well were all those parts of trees, even revenue-bearing trees, which might have been useful to the population but whose value could not be converted into fiscal receipts. Here I have in mind foliage and its uses as fodder and thatch; fruits, as food for people and domestic animals; twigs and branches, as bedding, fencing, hop poles, and kindling; bark and roots, for making medicines and for tanning; sap, for making resins; and so forth. (Scott, 1998: 12)

Scott sees such ‘tunnel vision’ as essential for governing complex phenomena. By focusing on some selected aspects only, they can be brought into sharp focus and made more legible, measurable and amenable for calculation. In education, the focus of such tunnel vision has been on testing of literacy and numeracy, and performance in literacy and numeracy assessments stands proxy for the efficacy and quality of a school system. Literacy and numeracy happen to be the most easily testable subjects on a large scale, particularly in international comparison, as one PISA official explained:

Reading, science and maths are there largely because we can do it. We can build a common set of things that are valued across the countries and we have the technology for assessing them. So there are other things like problem solving or civics and citizenship – that kind of thing where there would just be so much more difficulty in developing agreement about what should be assessed. And then there are other things like teamwork and things like that. I just don’t know how you’d assess them in any kind of standardized way So you are reduced to things that can be assessed. They’ve tried writing – but . . . the cross-cultural language effect seems too big to be comparable. So the things we assess are a combination of the things we value and the things we can do – I think it sends an odd message about science, perhaps, but I don’t think anyone would argue about literacy and numeracy. (Interview transcript, PISA expert, cited in Gorur, 2011)

It is an aphorism that we don’t just measure what we value, but that we come to value what we measure. The narrow focus on test outcomes has led not only to a narrowing of the curriculum (Berliner, 2011) in terms of education practice, but also to ‘governing by numbers’ (Rose, 1991) in terms of education administration. Drawing from Sahlberg (2015), Morris (2016) says that the OECD, through PISA, promotes ‘[t]he ‘standardization’ of education to focus on outcomes and testing’ and ‘[t]he elevation of core subjects such as literacy and numeracy that are more suited for measurement by global assessment surveys such as PISA’. Sahlberg (2011) coined the term GERM (Global Educational Reform Movement, of which PISA is a major part), to emphasise that, globally, education reform appears in the grip of a contagion – one that promotes competition, standardisation, test-based accountability and school choice in the service of a frenzied scramble to raise test scores and rankings. Other values and strategies such as collaboration, personalisation, trust-based professionalism and equity of outcomes are neglected. Just as the focus on timber led to the neglect, and eventual loss, of foliage, fruit, twigs and branches, bark, roots, sap and so on, the narrow focus on outcomes in terms of ‘literacies’ of a particular kind ignores other aspects of schooling that are arguably of importance to children and their parents, and indeed to society.

Losing the detail: from complex to abstract tree

The diversity of plant life in the forest was a nuisance; it presented a problem for measuring and calculating timber yield. The timber yields of different trees were not uniform, nor did the trees take the same length of time to grow and mature, or require the same soil conditions to flourish. This diversity needed to be tamed to make measurement and prediction more accurate. The Germans created five standardised categories of tree sizes, each assigned an anticipated timber

yield. Projecting from sample forest plots, they were then able to estimate the timber yield of the entire forest. Real trees were thus translated to abstract ones, with each tree belonging to one of five standard categories.

Comparison across diverse school systems and economies, which included countries with vastly differing cultures, goals, ambitions and social, political and economic situations was similarly problematic for PISA in its bid to develop internationally comparable accounts. To make such comparisons possible, a series of moves was required – students, test items and testing and scoring processes had to be strictly standardised and abstracted in order to render them comparable. Such abstraction works on many levels in PISA. The three types of ‘literacies’ it examines – reading, mathematical and scientific literacy – stand for ‘outcomes of the education system’. Within each domain of testing, individual test items are standardised so that they perform in the same way internationally. The item difficulty is standardised to aid calculations. Performance is benchmarked into six levels, with Level 1 being the most basic and Level 6 the highest. The 15-year-olds who participate in the survey are also standardised – the individual student, in all his or her complexity, is lost. Gone are the complex anxieties and excitements, and the goals and dreams and motivations and interests of 15-year-olds. In their place we have students defined economically by such attributes as their gender, by a proxy for a socio-economic level, by their nationality and their status as a migrant. We have a standardised student who simply becomes part of a yield or outcome measure. In fact, each PISA test is distributed among several students – so several students are grouped together to form an abstract student in PISA (see Gorur, 2011 for a detailed account).

The great advantage of such standardisation is that it renders the test easily adaptable to different scenarios, and thus to expansion. With the expectation now of expanding into more than 100 nations (PISA officials, interview transcript), PISA is entering the space of middle- and low-income nations with a modified version called PISA for Development (PISA-D). The pilot study is being planned in Ecuador, Honduras, Paraguay, Senegal, Guatemala, Zambia and Cambodia in 2016, with Panama being negotiated at the time of this writing. For these nations, where more students are expected to perform poorly, PISA is being modified to include larger numbers of Level 1 and Level 2 questions. However, the tests are being standardised in such a way as to allow students from PISA-D countries to be compared with their counterparts in PISA. To make these assessments comparable across such a diversity of contexts, detail has to be sacrificed.

Developing a synoptic view (education ‘at a glance’)

For German forestry to move from crude estimations and little control to scientific practices and efficient management, the forest had first to be made legible. The narrowing of the field of vision to timber, to the exclusion of other aspects of the forest, and the standardisation of trees, allowed a radical simplification, which in turn facilitated legibility – i.e., making nature and society readable and mappable – which, Scott asserts, is ‘a central problem in statecraft’ (Scott, 1998: 2). The standardisation of trees prepared the way for such legibility:

By radically narrowing his vision to commercial wood, the state forester had, with his tables, paradoxically achieved a synoptic view of the entire forest. This restriction of focus reflected in the tables was in fact the only way in which the whole forest could be taken in by a single optic. Reference to these tables coupled with field tests allowed the forester to estimate closely the inventory, growth, and yield of a given forest. (Scott, 1998: 15)

Mathematics was brought to bear in order to calculate the volume of saleable wood in a standardised tree of a particular age, in conjunction with the years needed for the tree to mature. Working

with actual timber yields from standardised trees, these calculations became quite sophisticated and helped in both managing forests (deciding when to fell specific trees, for example) as well as in predicting annual yields. It also served to set benchmarks, so that forests that were performing below par could be identified and interventions introduced.

By detaching children from around the world from their contexts, standardising them and converting them into numbers, the OECD is able to create sophisticated calculations, identify problems, and suggest solutions and policy advice with extreme specificity. Performance can be disaggregated on the basis of gender, migration status, social capital, location and other dimensions. Specific areas for intervention can thus be isolated. With each round of the survey producing more information, trend data create patterns of growth and decline. This is the type of calculus that ‘centres of calculation’ (Latour, 1987) can perform from afar, sitting in a distant office, with the numbers providing a synoptic overview of the entire phenomenon, if at the expense of detail:

An enormous variety of things have now been converted into inscriptions on a completed form, coded according to pre-designated structures, and brought safely to the PISA offices. The students and their learning—indeed, whole school systems, represented by these students, have all been ‘detached, separated, preserved, classified, and tagged’ (Latour, 1999: 39). The world has now been transformed into ‘two-dimensional, superposable, combinable inscriptions’ (Latour, 1999: 29), so that scientists are able to sit in the comfort of their offices and reassemble, reunite and redistribute them ‘according to entirely new principles that depend on the researcher, on the discipline ... and according to the institution that shelters them’. (Gorur, 2011: 88)

The value of a synoptic view is that it is available ‘at a glance’ – providing easily absorbed and easily represented information. PISA’s league tables, on which 15-year-old children from distant and diverse parts of the world are all gathered and organised into obedient rows and columns on a single spatio-temporal frame (Gorur, 2011), are a perfect example of such a synoptic view.

Creating standardised forests

The most astonishing step in German forestry management, however, was not that of making forests legible and rendering them into simplified maps, but in actually *creating* simplified, standardised forests that resembled the abstracted accounts of the forests in their ledgers:

What is decisive for our purposes, however, was the next logical step in forest management. That step was to attempt to create, through careful seeding, planting, and cutting, a forest that was easier for state foresters to count, manipulate, measure, and assess. The fact is that forest science and geometry, backed by state power, had the capacity to transform the real, diverse, and chaotic old-growth forest into a new, more uniform forest that closely resembled the administrative grid of its techniques. (Scott, 1998: 15)

Clearing the underbrush, reducing the number of species in a forest, and planting same-species, same-age trees in neat rows, German foresters began to create geometrically regular forests which were much more readily monitored and managed. The natural forest had been domesticated – no longer was it wild and disorderly. It was replaced by regimented, orderly rows of trees.

The forest trees were drawn up into serried, uniform ranks, as it were, to be measured, counted off, felled, and replaced by a new rank and file of lookalike conscripts. As an army, it was also designed hierarchically from above to fulfill a unique purpose and to be at the disposition of a single commander. At the limit, the forest itself would not even have to be seen; it could be ‘read’ accurately from the tables and maps in the forester’s office. (Scott, 1998: 15)

This translation of the abstract forests of the fiscal ledger books into real forests that mimicked the reductionist order of the ledger books epitomises the performativity of measurement and mapping. Here I use the term ‘performativity’ in the sense it is used in actor–network theory and science and technology studies (STS) more generally (for example, Law, 2009; Knorr-Cetina, 1999; Pickering, 1993), which posits that exercises such as measurements are not just representations or descriptions of existing realities, but participate in their production.

PISA, too, is bringing into being a kind of ‘standardised school system’ with its calculations and policy guidance and recommendations. PISA was introduced to provide some comparative measures of the outcomes of schooling in the OECD nations. However, PISA does much more than merely provide a snapshot of how countries perform in PISA. With each round of PISA, it produces not only the league tables with countries ranked according to performance, but also detailed analyses which correlate practices of schooling, funding patterns and policies with performance. It provides pointers to countries on how to improve their school systems. In addition to the PISA reports, the OECD uses PISA data to produce thematic reports on particular topics and country reports focusing on particular countries and how they could improve their systems, and also offers consultancy where OECD experts visit client nations to help reform their systems, bringing them closer in line with PISA ideals.

Because it is conducted by a transnational agency, PISA is seen as disinterested and objective, and therefore trustworthy. Countries have more faith in PISA’s judgement than they do in their own. Many countries were surprised by the findings of the first PISA results that appeared in 2001. Germany went into ‘PISA Shock’ at its poor showing relative to other OECD nations (Ertl, 2006; Gruber, 2006). PISA provided new role models, displacing old ones. Canada, which had been looking to Germany for policy lessons, learned that Canada’s own education system was in fact superior to Germany’s, as one interviewee pointed out to me:

Canadians used to be constantly going to Germany to study them so we could copy their system! I went to this meeting in Berlin in 2001, where their Federal Minister got up and said, ‘Well, we need to learn from Canada, because they’re doing so much better’, and I wanted to yell out, ‘Give us our money back for all the trips we’ve made!’ Finland has got new hotels to accommodate all the PISA visitors. And they were looking to Germany before the PISA results came out. Everyone was going to the US all the time – no one goes to the US any more to see how they do schooling ... no one thinks that the US is the international model for how to do schooling. (2008 Interview transcript: senior policy bureaucrat)

Similarly, Australia was evaluated as being ‘high quality, low equity’, because there was a greater correlation between socio-economic status and performance than was the case on average for OECD nations. This was very surprising to Australia. However, when there is a gap between a country’s perception of its system and PISA’s evaluation of it, it is PISA that is believed (Martens and Niemann, 2010). In other words, once PISA came along, countries began to ‘see like PISA’. PISA appears to displace easily previously held understandings and impose its own numeric imprint.

Trust in PISA has meant that many countries have reformed their systems in accordance with PISA’s recommendations. In an evaluation of the influence of PISA by Breakspear (2012), 28 participating countries reported that PISA had been very or moderately influential in informing policy at the national level. Many were able to point specifically to aspects of policy that had been influenced by PISA. A country’s rank – especially if it has ‘declined’, appears to spark policy debate and instigate policy changes. It would not be an exaggeration to say that the rankings continue to play a significant role in shaping policy debate and in influencing policy makers. PISA rankings have cemented the notion of education as a ‘race’ to be won – one that can only be won by

displacing others. It is a metaphor that has captured – and limited – the policy imagination. Similar views about the ‘education race’ are evidenced in the US reform introduced by President Obama called ‘Race to the Top’; in Australia’s stated ambition in its Education Act of being in the ‘top five’; and in the panic in some countries about ‘slipping’ and being ‘overtaken’ in the rankings.

A large survey conducted by Hopkins et al. (2008) found that a majority of the key stakeholders surveyed believed that the learning skills addressed in PISA were more aligned to their country’s policy needs than the skills assessed in their national tests and that, therefore, the skills valued in PISA should be more systematically addressed in school curricula. PISA thus influences the very goals of schooling, as well as schools’ curricular programmes. Respondents in the Hopkins et al. survey listed the following as the reforms most likely to be adopted:

- The development of national standards;
- The establishment of national institutes of evaluation;
- Changes in the curriculum or introduction of national curricula;
- The introduction of educational programmes targeting specific groups of students;
- An increase in the allocation of resources to schools; and
- An increase in collaboration amongst the key stakeholders of the education system within and across countries.

The survey reported that the changes in school policies and practices that had been adopted in some countries, based on PISA results, included increased autonomy for schools; the establishment or further development of accountability systems; discussion on increasing the number of hours spent in school; the introduction of selection examinations; a focus on the development of life skills and those valued by the labour market; the development of an interest in empirical educational research; and more fine-grained data analyses, in general, (e.g., focus on gender differences or the influence of socio-economic background).

Baird et al. (2011) conducted a study of six case-study nations to evaluate the impact of PISA. They found that PISA affected different countries in different ways, and included such responses as greater convergence in policy and increased standardisation across provinces in Canada, and ‘PISA shock’ in Norway, which the government exploited to bring in a spate of reforms to curriculum and assessment. OECD itself reports on its influence, saying that PISA has become ‘the world’s premier yardstick for evaluating education systems’:

Over the past decade, the OECD Programme for International Student Assessment (PISA), has become the world’s premier yardstick for evaluating the quality, equity and efficiency of school systems in providing young people with these skills. But the evidence base that PISA has produced goes well beyond statistical benchmarking. By identifying the characteristics of high-performing education systems, PISA allows governments and educators to identify effective policies that they can then adapt to their local contexts. (OECD, 2014: 2)

Despite reservations and caveats, ‘learning from high-performing nations’ has been a significant part of PISA’s influence. A major area of influence is in the development and use of national assessments and evaluation systems. Countries also reported changes to curricula to emphasise ‘PISA-like competencies’; the adoption of particular strategies to improve reading literacy performance; changes to financing decisions to enhance equity; and strategies to monitor the performance of ‘top’ or ‘bottom’ performers (Breakspear, 2012: 18). The OECD produces reports that specifically engage with what individual countries can learn from PISA – for example, *Lessons from PISA for Japan: Strong Performers and Successful Reformers in Education* (OECD, 2012), and *Lessons*

from PISA for the United States: Strong Performers and Successful Reformers in Education (OECD, 2011). Indeed, PISA is creating new ‘reference societies’ from which policies are being borrowed, and against which countries are evaluating themselves (Waldow, 2015). A statement by Michael Gove, then UK Secretary of State for Education,² epitomises the compelling force of PISA’s logic and evidence:

No nation that is serious about ensuring its children enjoy an education can afford to ignore the PISA and McKinsey studies [*How the world’s best-performing schools come out on top* (Barber and Mourshed, 2007)]. Doing so would be as foolish as dismissing what control trials tell us in medicine. It means flying in the face of the best evidence we have of what works ... our recently published Schools White Paper was deliberately designed to bring together – indeed, to shamelessly plunder from – policies that have worked in other high-performing nations. (Gove, 2012, cited in Morris, 2016)

By ‘flattening’ education into a standardised, decontextualised phenomenon, education *à la* PISA has become portable. Its ideas and philosophies are seen as being readily and easily transportable across the world, so that lessons from Shanghai and Singapore can be implemented in Azerbaijan or Peru. As detailed above, PISA has influenced the very fundamentals of education – curricula and assessment. Critically, it has also facilitated and encouraged particular forms of accountability and control, as I discuss below.

Governing by template

The value of legibility and standardisation for centralised control cannot be over-emphasised (Porter, 1995). One of the great affordances of the regimented scientific forests was that they were far easier to manage. With hardly any shrubs or grasses, clearing the underbrush became much easier: so did felling and planting. Indeed, the regularity of the forest facilitated a certain routinisation and standardisation of forestry practices which could be specified, codified and inscribed into training protocols. Using these protocols, relatively unskilled workers could perform the required tasks in the new forest environment. The uniformity of the same age, same species forests made the predicting of timber yields much more reliable. This in turn helped in marketing the timber.

The new standardised forest also became a veritable laboratory, where, with diversity drastically minimised and many variables held constant, experiments to examine the effects of ‘fertilizer applications, rainfall, and weeding, on same-age, single-species stands’ (Scott, 1998: 18) became feasible. The careful calculations enabled by the highly legible forests pointed towards some ‘winners’ – trees that were the most productive and cost-efficient to grow. Certain species, such as the hardy and rapid-growing Norway spruce, became a favoured tree, yielding stunning commercial results. The new scientific forest also became a powerful aesthetic, its regularity and neatness greatly valued. Underbrush and fallen branches were regularly removed to keep the forest floor clean.

Similarly, with its gaze focused on a narrow set of actors and a narrow scope of action, PISA is able to develop ‘best practices’ to improve the efficiency of school systems. ‘Winners’ are identified and studied, the better to emulate them. Publications such as *How the World’s Best-Performing Schools Come Out On Top* (Barber and Mourshed, 2007); *Lessons from PISA for the United States: Strong Performers and Successful Reformers* (OECD, 2011); *What We Learn from the PISA 2012 Results* (Schleicher, 2013); and *Catching Up: Learning from the Best School Systems in East Asia* (Jensen, 2012), to name but a few, have come to be produced. Indeed, the OECD regularly produces thematic issues and other publications with policy lessons based on PISA. That many such

publications are produced by management consulting companies is not a trivial detail – in the new, standardised school systems of PISA calculations, it is possible to (mis)understand teaching and school administration as standardisable and routinisable practices which can be codified in universally applicable terms – as matters of management rather than pedagogic expertise.

PISA has also contributed to practices of accountability and transparency through the generation and use of numeric data to facilitate the imposition of control from afar (Porter, 1995; Scott, 1998) and, in general, a convergence of policy (Halpin and Troyna, 1995). A new normativity – a morality, even – has come to be imposed by PISA, and absorbed by PISA participants. This is not just about adopting a specific policy, such as teaching phonics; it is about putting in place an ongoing system of ever increasing monitoring and data generation, and about fundamentally changing the structures and systems of administering and governing schools. It is about changing the very culture of a society by influencing the curriculum (Zajda, 2001).

Seeing like PISA

The phenomenon of ‘seeing like PISA’ is not just about the influence of PISA on national policies. Rather, it is about a particular set of approaches and understandings that are epitomised by PISA. These approaches and practices were already being actively promoted globally even before PISA was developed. ‘Seeing like PISA’ – i.e., standardisation, the development of a narrow field of vision focused on literacy and numeracy outcomes, abstraction, and the generation of standardised templates and protocols to guide practices – have been proliferating since the 1980s, as have the associated governance and administrative approaches of performance-based accountability with rewards and punishments, continuous monitoring and market-based incentives. The global narrative of education performance as a ‘race’ to be won, and the regular comparisons provided by PISA and similar assessments, have intensified into a frenzied approach to reform. Andy Hargreaves (2011: xvi) suggested that the current reform climate in education reflects a culture of ‘an unhealthy obsession with all that is bigger, harder, tougher, faster, and stronger’, resulting in unsustainable pressures and long-term disasters, such as ecological damage inflicted by companies chasing financial gain, and a global financial crisis resulting from unrealistic targets and greed. He cites the current ‘school reform on steroids’ as characterised by ‘failure, firings, competition, and closures’.

‘Seeing like PISA’, is characterised by a reliance on numbers; enhanced mechanisms of accountability; a heightened focus on education as an economic commodity; a proliferation of testing; and the viewing of education as a global race with winners and losers. The shift in focus to outcomes rather than inputs, the promotion of test-based accountability systems, and the promotion of corporate models of management, tied to a focus on teacher quality and schemes to enhance the surveillance and improvement of teachers, have all become nearly universal. PISA’s trust in market-based approaches, its focus on economic gains and returns, its views on evidence, its faith in surveys and psychometrics, its focus on competition as a driver of quality and the governance models it promotes, based on heightened forms of accountability and surveillance are the hallmarks of our times, and are pursued in a range of other social policy arenas.

PISA is a prime example of – as well as an active participant in – GERM (Sahlberg, 2011). Sahlberg elaborated the phenomena that contributed to GERM: a new paradigm of learning inspired by constructivist approaches that shifted attention from teaching to learning; public demand for ‘guaranteed, effective learning for all pupils’ and common standards, leading to centralised curricula and centralised assessment programmes; and the decentralisation of services and devolution of responsibility with the accompanying proliferation of accountability practices. Sahlberg identified six unintended negative consequences of GERM: standardisation and a focus on outcomes leading to a belief among policy makers that ‘setting clear and sufficiently high performance

standards for schools, teachers, and students will necessarily improve the quality of the desired outcomes'; an increased focus on literacy and numeracy, which are seen as core subjects; teaching toward pre-determined results using 'guaranteed content' and 'proven methods', limiting teacher autonomy and experimentation – and thus leading to teacher deprofessionalisation and the 'McDonaldisation' of teaching; the imposition of ideas from the corporate world to foster reform, resulting in the loss of opportunity to approach reform using education's own past experiences in an organic fashion; test-based accountability practices that are linked with 'processes of accrediting, promoting, inspecting, and ultimately, rewarding or punishing schools and teachers'; and market-based approaches to schools premised on school choice, and the consequent increase in control over schools through centrally mandated standards, a tightening of administrative control over teachers, and an increased focus on data from students' performance in standardised tests, in turn resulting in teachers and schools going to extreme lengths to secure better results for their students on these 'high stakes' tests (Sahlberg, 2011: 178-179)

These aspects of 'seeing like PISA' are heavily interrelated. Standardisation aids the translation of complex issues into simply stated problems that suggest simplistic solutions which are measurable and can be tracked over time. Standardisation and a focus on literacy and numeracy allows comparisons to be made, in turn creating competition and particular forms of surveillance of teachers and schools. The pressures applied by this result in the deprofessionalisation of teachers, aided by decontextualised, abstract understandings of teaching and learning, and indeed of education administration and governance.

It is the interrelated nature of these practices, and their embeddedness in multiple networks, that makes the destabilisation or undoing of this network so difficult (Latour, 1987). An example from PISA itself serves to elaborate this idea. Because PISA aims to collect data that are comparable over time in order to establish trends, its ability to change the tests to reflect advances in testing science or new developments in language (for example, acknowledging the language of texting and twitter as styles of communication) is highly restricted – doing so would compromise the trend comparisons. The enormous investment made originally to develop PISA gets in the way of change, even if changes might actually provide better assessment of literacy. Meanwhile, schools and school systems may be continuing to reform their curricula in line with the types of questions asked in PISA. The same types of questions might become embedded in national assessments. When new initiatives are developed, such as PISA-D, these protocols are embedded in even more networks. After a point, undoing any aspect of such networks becomes too costly to warrant the change.

Even as PISA has continued to expand and grow ever more popular, there has been alarm and dismay at the influence it is having on national policies and its apparent ability to play 'Pied Piper' and attract a vast and diverse range of nations to heed its directives. So concerned were some academics, education practitioners and public intellectuals that they decided to petition the OECD by writing a letter³ to Andreas Schleicher, Director of PISA, in May 2014, outlining a range of concerns about the use of PISA and suggesting some solutions. Initiated by US academic Heinz-Dieter Meyer and a New York Principal, Katie Zahedi, the letter was signed by prominent academics like Stephen Ball, David Berliner, Henry Giroux and Diane Ravitch, and created a brief stir as it circulated through emails and was reported in the press. The issues it raised included:

- An escalation in standardised testing at national levels encouraged by PISA, and the increasing policy reliance on such thin, quantitative accounts;
- A shift towards short-term fixes to climb the PISA ladder, in favour of thoughtful strategies for sustained and long-term benefit, driven by the three-year cycle of PISA assessments;

- An overzealous focus on tested literacies to the detriment of unmeasured subjects and benefits of schooling, such as moral, physical and social development of students;
- An instrumentalist focus on employment as the main desired outcome of schooling
- Unequal ability to influence PISA, with only the OECD member nations having any power to influence the tests; and
- A blurring of boundaries as OECD embraces partnership with commercial vendors (Meyer et al., 2014)

The letter goes on to add:

... most importantly: the new PISA regime, with its continuous cycle of global testing, harms our children and impoverishes our classrooms, as it inevitably involves more and longer batteries of multiple-choice testing, more scripted “vendor”-made lessons, and less autonomy for our teachers. In this way PISA has further increased the already high stress-level in our schools, which endangers the well-being of our students and teachers. (Meyer et al., 2014)

Despite the weight of reputed public thinkers and its direct access to Andreas Schleicher at the OECD, the movement to slow or reform PISA had little influence on the OECD or on its PISA plans. Far from slowing down ‘the testing juggernaut’ as the letter advised, PISA is continuing to expand its reach.

Beyond PISA, there has also been considerable criticism of ‘seeing like PISA’ and the practices such approaches are fostering. There is alarm that commercial and philanthropic organisations are gaining increasing influence on schooling (Lingard and Sellar, 2014; Ball, 2009). Increasingly, education is being outsourced through such initiatives as Charter Schools. Cookie-cutter understandings and approaches to both teaching and education governance have now become pervasive. There is a flattening of spaces as curricula and assessment are going global, and spreading, in particular, to the global south. The UNESCO Institute of Statistics and the Brookings Institution have collaborated to develop a framework for *universal* learning embracing seven domains of learning in the pre-secondary years of schooling. They are now looking to develop assessment systems with indicators in each of the domains to facilitate ‘*global tracking*’ (UNESCO-UIS and Brookings Institution, 2013, p. 11, emphases added).

A disaster over time? Learning from the parable

The scientific forestry practices of Germany were initially extremely successful. Timber yield increased and the quality of timber improved. The rotation time was reduced and, overall, there was a much better economic return from the land. The German practices became widely admired and spread to other parts of the world and eventually became hegemonic, codified and taught in universities. The simplification of the German forests made possible the technicisation of the practices of forestry: it was this simplified practice that came to be taught and used widely.

It takes about 80 years for trees to mature and a crop rotation to occur. It was only with the second planting, almost a century later, that the negative effects began to be recognised. The second crop of conifers did not grow well. Production dropped by 20–30%. The reasons for this decline were complex. The nutrient cycle had been severely disrupted by the changes in the relationship between ‘fungi, insects, mammals and flora’ (Scott, 1998: 20) that participate in the processes of soil building and nutrient uptake. The fastidiousness of the Germans in keeping the forest floor clean, and the resultant ‘absence of woody biomass’ led to a reduction in the diversity

of fauna – in turn inhibiting the soil-building processes. Nutrients in the soil were used up and not replenished. Reducing the diversity in species, often to mono-species planting, meant that the entire forest was vulnerable to diseases or pests that favoured that species. Addressing these issues required large sums of money to be spent on fertilizers and pesticides (with, one imagines, further damage to the environment, including contamination of the soil and water). The very factors that were bracketed out during the calculations came back to haunt the forest managers. The initial success was in large part due to the ‘soil capital’ that was present in the forests – and this had been depleted by the scientific forestry practices. To address this situation, the Germans invented a science called ‘forest hygiene’ to identify and replace what was missing in the scientific forest:

In place of hollow trees that had been home to woodpeckers, owls, and other tree-nesting birds, the foresters provided specially designed boxes. Ant colonies were artificially raised and implanted in the forest, their nests tended by local schoolchildren. Several species of spiders, which had disappeared from the mono-cropped forest, were reintroduced. (Scott, 1998: 21)

Interestingly, rather than reverting to diverse forests, the remedy devised was in the form of compensation, whilst the essential condition of lack of diversity remained. As Scott notes, the attempt was to create a ‘*virtual* ecology, while denying its chief sustaining condition: diversity’ (Scott, 1998, his emphasis).

The lesson that Scott draws from this parable is best expressed in Scott’s own words:

The metaphorical value of this brief account of scientific production forestry is that it illustrates the dangers of dismembering an exceptionally complex and poorly understood set of relations and processes in order to isolate a single element of instrumental value. (Scott, 1998: 21)

In education, statistical regression analyses such as those employed by PISA, which ‘control for’ various factors to develop universal ‘policy lessons’ are an example of such practices. Such aspects as gender, ethnicity and socio-economic status, for example, can be segregated in analyses, but they participate in very complex and interrelated ways in real classrooms, where their presence cannot be bracketed out. Many of these phenomena are relationally produced – and ‘isolating’ them as separate variables represents a fundamental ontological misunderstanding. Furthermore, although statistical analyses have become considerably sophisticated and can accommodate the notion of relationality, the calculations are premised on models constructed from inadequate understandings of the complexity of societies and classrooms – or at least, the models are not able to reflect these complexities. In any case, detailed analyses, even when available, are seldom used to inform policy, given the compelling lure of the rankings and league tables (Gorur, 2011; Gorur and Wu, 2015).

PISA’s ‘utilitarian simplification’ – its single-minded focus on ‘outcomes’ and ‘student performance’ on a limited range of literacies, the valued ‘commodities’ of education – may lead to increased ‘productivity’ in the short term, though even this claim is a stretch. Many countries have imposed a series of policy changes only to see their country’s scores decline – Australia is a very good example of this (Thomson et al., 2013). Few nations have shown any steady increase even in the limited range of PISA’s ‘outcome’ measures.

The teaching of a range of other aspects of education – culture and values, respect, aesthetics, empathy and compassion – may gradually be reduced given the utilitarian focus on ‘literacy and numeracy’ and the skills valued by the labour force. This ‘narrowing of the curriculum’ is already being widely reported and lamented (Crocco and Costigan, 2007). The effect of a curriculum where such features are given minimal importance and only certain utilitarian skills emphasised is difficult to predict, but arguably is likely to be significant.

PISA it is trying to address the issue of the narrowing of focus to reading, mathematical and scientific literacy by attempting to expand the range of what it tests. PISA introduced assessment in financial literacy in 2012. It is currently working on developing assessments of 'global competence' to assess 'young people's attitudes, values and knowledge of global issues', according to the OECD website,⁴ which adds that its assessment of global competencies would offer the first, comprehensive overview of education systems' success in equipping young people to support the development of peaceful, diverse communities'. It proposes to extend the 2018 assessment to include the assessment of global competencies. The re-introduction of such values after first depleting education through the narrow focus on the literacies valued for employment and productivity resonates with the deliberate and artificial introduction of diversity into mono-cropped forests.

The simplification of forests (or education systems) in order to measure them and make them legible, Scott points out, is not remarkable on its own – it is only through such tunnel vision that statecraft is possible. However, he asserts, when that simplification and legibility is accompanied by a high-modernist ideology, a passive citizenry and a time (or, I add, a *narrative*) of crisis, such simplification can lead to practices that could ultimately prove dangerous. The OECD accompanies the radical simplification of PISA with a high-modernist ideology in which increasingly sophisticated measurement in all spheres of social life will lead to better and better policy-making, resulting in 'the economic and social well-being of people around the world'. Given their economic bias, their very first aim is to 'restore confidence in markets and the institutions that make them function'.⁵ Because the OECD is a club joined voluntarily by nations actively *seeking* policy advice from it, the 'compliant citizenry' aspect is easily satisfied. Finally, a narrative of crisis constructed around an uncertain future of economic upheavals, aided by the continuing effects and memories of the recent global financial crisis, and now the increase in terrorism and involuntary migration, provide the conditions for the worst effects of 'seeing like PISA' to come into play.

The conditions, then, are ripe for the narrow and reductionist practices of 'seeing like PISA' becoming a disaster over time. PISA has been around for a relatively short period – a mere 16 years at the time of this writing. The first students to take PISA tests are just around 30 years in age – still early in their careers. As such, we are yet to see their influence on the world, and so the magnitude of the changes caused by the policy measures triggered by PISA may not be apparent for a considerable time. It would take many years before the results of such changes come to be manifested in any statistically significant way in assessments, and these assessments are, in any case, very narrow in scope. It will take even longer for the current cohorts of students to be of an age where their influence on the world can be felt in a pervasive way. In the meanwhile, the major overhauls of curricula, the introduction of 'high-stakes testing', the attendant corruption that is widely reported (see, for example, Nichols and Berliner, 2007), the effects on the teaching workforce, and changes in the wider ecology of students' lives might all become costly problems to address. The 'ecology' that builds the school's intellectual and moral character may become depleted as the whole school focuses on 'data' and on raising the 'yield' in terms of particular skills.

If the cocktail of a narrow vision, widespread standardisation and abstraction, an exclusively fiscal view, a depleted curriculum, deprofessionalised teachers and market-driven accountability systems which are currently in evidence (and are expanding into the global south) continues unchecked, we can only speculate on the effects this will have not only on the economy, but also on the moral, intellectual and ethical fibre of society. Currently, significant events such as global warming and the refugee crisis are testing our ability to empathise, our resourcefulness and our generosity globally. A market and choice-oriented approach to education might not prepare us well enough to respond to such crises in the best or most sustainable ways.

Two current examples illustrate the extent of standardisation and abstraction that has occurred in education policies and practices, and the performativity of such abstractions. The first example

is that of the US-based commercial company Bridge International Academies, which has already set up schools in several African nations, and to which Liberia has recently outsourced its entire pre-primary and primary education system. Bridge describes itself as follows:

We are data-driven and technology-enabled. Using smartphones and tablets, our 'closed loop' Learning Lab enables us to monitor teacher and student performance in real time, constantly reviewing and revising to ensure that we are offering a world class education that will prepare our students for the 21st century. (Bridge International Academies, 2016)

Through what has been dubbed 'school-in-a-box', Bridge is able to provide affordable schooling by eliminating teachers from its schools. Instead, highly scripted lessons developed by distant experts are delivered by individuals without teacher qualifications to large classrooms. In this way, cost is kept very low. Bridge aims to educate ten million students by 2025⁶ using this 'school-in-a-box' model.

The second example is that of a mathematical formula developed by a team of University of Melbourne economists at the Melbourne Institute of Applied Economic and Social Research to calculate the effectiveness of Principals. They claim that '[m]ore effective principals can raise student performance by as much as 0.22 of a year of learning' (The Melbourne Newsroom, 2016). This is an astonishing calculation, for several reasons. It is difficult to imagine a standardised measure of student performance being so accurate as to calculate difference in learning between schools in terms of 0.22 of a year. Further, causality even between teachers and their students' performance is impossible to establish statistically; to connect Principals to students' test performance would be a very long bow to draw. Nevertheless, the Minister of Education of the State of Victoria, in which Melbourne is located, has already endorsed this 'research' and plans to use it in training and supporting Principals. It would be hardly surprising if this calculation featured in the appraisal of Principals in course of time. It is through such processes that reductionist calculations come to be mimicked and produced as realities.

Scott's example of German scientific forestry management practices serves as a way to understand the steps involved in making phenomena legible through a narrowing of vision, standardisation and abstraction. The picture he paints of the gradual transformation of forests with all their diversity of flora and fauna into the standardised mono-cultural forests standing in tidy rows and columns provides a compelling image for us to stop and consider the possibility of the irreversible damage we may cause if 'seeing like PISA' goes unchecked. It is one thing to use such maps to aid governance, but quite another when we pave the way for commercial companies to proliferate 'schools-in-a-box', particularly to already vulnerable populations in the global south. If it is too soon to tell whether the effects of 'seeing like PISA' are dangerous or damaging, it is also too soon to have unmitigated faith in such an approach.

Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. <https://www.oecd.org/about/>

2. Note that the Department for Education, for which Michael Gove was the Secretary of State, attends to educational policy for England: the devolved governments for Scotland, Wales and Northern Ireland have their own Departments and policies.
3. <http://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics>
4. <https://www.oecd.org/pisa/oecd-proposes-new-approach-to-assess-young-peoples-understanding-of-global-issues-and-attitudes-toward-cultural-diversity-and-tolerance.htm>
5. <http://www.oecd.org/about/>
6. <http://www.bridgeinternationalacademies.com/company/about/>

References

- Baird J-A, Isaacs T, Johnson S, Stobart G, Yu G, Sprague T and Daugherty R (2011) *Policy Effects of PISA*. Oxford: Oxford University Centre for Educational Assessment.
- Ball SJ (2009) Privatising education, privatising education policy, privatising educational research: network governance and the 'competition state'. *Journal of Education Policy* 24(1): 83–99.
- Barber M and Mourshed M (2007) *How the world's best-performing school systems come out on top*. McKinsey and Company.
- Berliner D (2011) Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education* 41(3): 287–302.
- Breakspear S (2012) *The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance*. OECD Education Working Paper (Volume 71). Paris: OECD.
- Bridge International Academies (2016) Bridge International Academies. Available at: <http://www.bridgeinternationalacademies.com/company/about/> (accessed 02 June 2016).
- Crocco MS and Costigan AT (2007) The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education* 42(6): 512–535.
- Gorur R (2011) ANT on the PISA Trail: Following the statistical pursuit of certainty. *Educational Philosophy and Theory* 43(S1): 76–93.
- Gorur R (2014) Producing calculable worlds: Education at a glance. *Discourse: Studies in the Cultural Politics of Education* 36(4): 578–595.
- Gorur R and Wu M (2015) Leaning too far? PISA, policy and Australia's 'top five' ambitions. *Discourse-Studies in the Cultural Politics of Education* 36(5): 647–664.
- Gove M (2012) Education for economic success: speech by Michael Gove to the Education World Forum, 11 January 2012. Available at: <https://www.gov.uk/government/speeches/michael-gove-to-the-education-world-forum> (accessed 17 June 2016).
- Halpin D and Troyna B (1995) The politics of education policy borrowing. *Comparative Education* 31(3): 303–310.
- Hargreaves A (2011) *Foreword: 'Unfinished Business', Finnish Lessons: What Can We Learn From Educational Change in Finland?* New York and London: Teachers College Press.
- Hopkins D, Pennock D, Ritzen J, Ahtaridou E and Zimmer K (2008) *External Evaluation of the Policy Impact of PISA*. Volume EDU/PISA/GB(2008)35/REV1. Paris: OECD.
- Jensen B (2012) *Catching Up: Learning From the Best School Systems in East Asia*. Melbourne: Grattan Institute.
- Knorr-Cetina K (1999) *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.
- Latour B (1987) *Science in Action – How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.
- Latour B (1996) *ARAMIS or The Love of Technology* (Porter C, transl.). Cambridge, MA and London: Harvard University Press.
- Latour B (1999) *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA and London: Harvard University Press.
- Law J (2002) *Aircraft Stories: Decentering the Object in Technoscience*. Durham, NC: Duke University Press.

- Law J (2004) *After Method: Mess in Social Science Research* (1st edition). London and New York: Routledge.
- Law J (2009) Actor network theory and material semiotics. In: Turner BS (ed.) *The New Blackwell Companion to Social Theory*. Wiley-Blackwell, Oxford, pp. 141–158.
- Lingard R and Sellar S (2014) Globalization, edu-business and network governance: the policy sociology of Stephen J Ball and rethinking education policy analysis. *London Review of Education* 11(3): 265–280.
- Martens K and Niemann D (2010) *Governance by comparison: How rating and ranking impact national policy making in education*. TranState Working Paper. Bremen: University of Bremen Collaborative Research Centre.
- Massumi B (2002) Introduction: Concrete is as concrete doesn't. In: Massumi B (ed.) *Parables for the Virtual: Movement, Affect, Sensation*. Durham, NC: Duke University Press, pp. 1–22.
- Meyer H-D, Zahedi K and Signatories (2014) *An Open Letter: To Andreas Schleicher, OECD, Paris*. Available from: <http://www.globalpolicyjournal.com/blog/05/05/2014/open-letter-andreas-schleicher-oecd-paris> (accessed 17 June 2016).
- Mol A. (1999) Ontological politics: A word and some questions. *The Sociological Review* 47(S1): 74–89.
- Morris P (2016) *Education Policy, Cross-national Tests of Pupil Achievement and the Pursuit of World-class Schooling: A Critical Analysis*. London: UCL Institute of Education Press.
- Nichols SL and Berliner DC (2007) *Collateral Damage: How High-Stakes Testing Corrupts America's Schools*. Cambridge, MA: Harvard Education Press.
- OECD (2010) *The High Cost of Low Educational Performance: The Long-run Economic Impact of Improving PISA Outcomes*. Paris: OECD.
- OECD (2011) *Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education*. Paris: OECD.
- OECD (2014) *PISA 2012 Results in Focus: What 15-Year-Olds Know and What They Can DO With What They Know*. Paris: OECD.
- OECD (2012) *Lessons from PISA for Japan: Strong Performers and Successful Reformers in Education*. Paris: OECD.
- Pickering A (1993) The mangle of practice: Agency and emergence in the sociology of science. *American Journal of Sociology* 99(3): 559–589.
- Porter TM (1995) *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- Rose N (1991) Governing by numbers: Figuring out democracy. *Accounting Organizations and Society* 16(7): 673–692.
- Sahlberg P (2011) *Finnish Lessons: What Can the World Learn from Educational Change in Finland*. New York and London: Teachers College Press.
- Sahlberg P. (2015) *Finnish Lessons 2: What Can the World Learn From Educational Change in Finland?* New York and London: Teachers College Press.
- Schleicher A (Producer) (2013) What we learn from the PISA 2012 results? *Education & Skills Today*. Available at: <http://oecdeducationtoday.blogspot.com.au/2013/12/what-we-learn-from-pisa-2012-results.html> (accessed 17 June 2016).
- Scott JC (1998) *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.
- Serres M and Latour B (1990) *Conversations on Science, Culture and Time* (Lapidus R, Transl.). Ann Arbor, MI: The University of Michigan Press.
- The Melbourne Newsroom (2016) School principals directly impact student test scores. Available at: <http://newsroom.melbourne.edu/news/school-principals-directly-impact-student-test-scores> (accessed 02 June 2016).
- Thomson S, De Bortoli L and Buckley S (2013) *PISA 2012: How Australia Measures Up*. Camberwell, Victoria: Australian Council for Educational Research Ltd.
- UNESCO Institute for Statistics and the Centre for Universal Education at Brookings (2013) *Toward Universal Learning: Recommendations from the Learning Metrics Task Force* (Volume Reference UIS/2013/ED/TD/07. Washington, DC: UIS/Brookings.

- Waldow F (Producer) (2015) Projecting images of the 'good' and the 'bad' school: Using top scorers in large-scale assessments as reference societies (Keynote video, *New York City Seminar: International Assessment and International Convergence of Educational Reforms*, Columbia University New York, 6–7 March). Available at: <http://international-assessments.org/seminar-2/> (accessed 17 June 2016).
- Woolgar S (1991) Beyond the citation debate: Towards a sociology of measurement technologies and their use in science policy. *Science and Public Policy* 18(5): 319–332.
- Zajda JI (2001) *Curriculum, Culture and Teaching*. Albert Park, Australia: James Nicholas Publishers Pty Ltd.

Author biography

Radhika Gorur is a senior lecturer at Deakin University, Australia, and a Director of the Laboratory for International Assessment Studies. The examination of contemporary practices of quantification in education has been central to her research agenda. Using material-semiotic approaches and concepts from Science and Technology Studies (STS), she has been developing a 'sociology of measurement and numbers' that makes explicit the instrumental and constitutive work of quantification, calculation and comparison in policy. She is currently studying initiatives in assessment and accountability in low-income nations and exploring the possibilities of inclusive, collective and sustainable accountability practices.