

Historicizing and contextualizing global policy discourses: Test- and standards-based accountabilities in education

Bob Lingard

The University of Queensland

This paper in commenting on the contributions to this special number demonstrates the necessity of historicizing and contextualizing the rise of test- and standards-based modes of accountability in contemporary education policy globally. Both are imperative for understanding specific national manifestations of what has become a globalized educational policy discourse, namely, that such modes of accountability will drive up student performance and thus enhance the global economic competitiveness of nations. New modes of testing might be seen as a new-old system, given the provenance of testing in schooling systems. The paper will argue that there are global and national elements to this policy situation, with national and global testing of students and school systems sitting in symbiotic relationships with each other. It will also be argued that high-stakes testing has become a fourth message system of schooling that steers today the practices of schools and teachers in classrooms, often with reductive anti-educational effects. Comment will also be made about how these educational developments fit within the broader structure of feeling and how they open up possibilities for edu-businesses to take an increasing role in policy.

Keywords: contemporary education policy, education policy discourse, student performance, performance testing, high-stakes testing

In this afterword I will write across the papers in the collection, attempting to draw out the issues they raise. I will also locate the collection of papers within debates about the reductive educational effects of high stakes testing and new educational accountabilities and how they fit within the contemporary policy assemblage. High-stakes tests are a central component of the neo-liberal policy settlement in contemporary education globally and also linked to new modes of governance in education (Rizvi & Lingard, 2010), as the papers so ably demonstrate. The latter has seen new public management restructure state bureaucracies associated with Keynesian, while more recent moves have witnessed the rise of network governance as well (Ball & Junemann, 2012). The discursive and governance modes are now played out globally as well as nationally

and locally, as I will argue, with implications for a politics of resistance to the neo-liberal policy settlement.

HISTORIZING AND CONTEXTUALIZING TEST-BASED ACCOUNTABILITIES

This collection of papers deals with vernacular expressions of a global education policy mantra that new forms of test- and/or standards-based accountability will improve school performance as measured by student performance, most often on tests of literacy and numeracy and sometimes science. Despite the globalized nature of this discourse, and what we might see, superficially at least, as policy convergence globally, what we have in reality, on the ground in systems, schools and classrooms, are vernacularized and hybridized manifestations of this discourse. The papers in this special number demonstrate unequivocally that context (e.g. geopolitical location, culture, politics, political structure, levels of inequality, place and structure of schooling systems, strength and focus of the teaching profession and teachers' unions) and history (e.g. historical role of schools in terms of opportunity structures, production and reproduction of elites, testing) matter. To reiterate – history and context matter, with context somewhat attenuated in these globalizing times with new imbrications and topological relations between the local, provincial, national, regional and global policy places and spaces.

This analytic necessity of acknowledging context and history is so in terms of the specific vernacular expressions of the globalized policy discourses, whether articulated as test- or standards-based accountabilities or both, and in terms of understanding why this particular expression or particular policy bricolage in these schooling systems and nations. This observation is perhaps best evident in John O'Neill's paper in this special number that documents New Zealand's take-up of standards-based rather than test-based accountabilities, a result of historic-cultural factors according to his analysis and resistance by the profession. It is also apparent in Thomas' account of the historical backdrop to more recent federal policy initiatives in respect of educational accountability in the USA, namely George W. Bush's *No Child Left Behind* and Barack Obama's *Race to the Top*. Indeed, Thomas traces the genealogy of these reforms back to the late nineteenth century and the emergence of scripted core curriculum and high-stakes testing. He also documents the historical backdrop to the development of a test-based approach in the US state of South Carolina. Wayne Au likewise traces a similar history of testing going back to what he refers to as a US appropriation of the IQ tests developed by the French psychologist, Alfred Binet, in the early twentieth century.

All of the papers also demonstrate, in relation to Bernstein's (1971) three message systems of schooling, namely curriculum, pedagogy and evaluation, that the latter has become *a*, if not *the*, major policy steering mechanism in all of the national cases dealt with. The evaluation message system, rearticulated as high stakes testing or standards-driven expressions of new educational accountabilities, has affected to a considerable degree both the curriculum and pedagogy in schools and classrooms (Nichols &

Berliner, 2007, Rizvi & Lingard, 2010, Lingard, Martino & Rezai-Rashti, 2013). As with Thomas in this special number, we also need to distinguish between standardised testing and teachers' formative assessment practices in classrooms. Howard Stevenson and Philip Wood in their paper document and demonstrate the impact of high stakes testing on teachers' work in England in an era of austerity and downward pressures on public expenditure. They show as well how teachers' labour has been converted into a product framed by various metrics linked to testing of various kinds. Some time ago now, Connell (1985) argued that school teaching was a labour process without a product. This observation no longer holds true. Greg Thompson analyses survey data about the impact of national literacy and numeracy testing in Australia on teachers' work as well and on their identities. His work confirms the international research literature on the impact of high stakes testing: narrowing of curriculum, teacher-centred pedagogy, teaching to the test, a reduction in student motivation and greater stress on students. We might even speak of high-stakes testing and standards-based accountabilities as the new fourth message system of schooling, which as would be argued by Bernstein (1971), sits in a symbiotic relationship with the other systems, but which is now driving them and is a central element of the contemporary policy assemblage in education: what might be seen as a new-old system, given the historical etymology of statistics as state numbers and the provenance of testing. The latter point is well made in an insightful fashion by Wayne Au in his paper in this special number. Stevenson and Wood make a point about how older modes of testing in England were linked to the sorting and sifting function of schooling, while the newer mode is in part about controlling teachers' work.

As noted in the Introduction to this paper, test-based accountability is part of a broader policy assemblage. Ball (2013) has documented three discourses and practices that frame this assemblage and the education policy work of nation states today. These are new managerialism (manifesting as new public management and more recently as network and heterarchical governance), choice and market reforms, and a new culture of performativity. These are assembled together in particular ways in different nations and change over time as well. As with Stevenson and Wood in this special number, I note though, that both test- and standards-based manifestations of educational accountability sit very comfortably with these other policy discourses, indeed are almost endemic to each of them. Stevenson and Wood argue that these discourses create teacher labour as product, with teachers caught in a pincer between marketisation and the new managerialism and measuring their labour outputs is central to both. I would argue they are central to new centre/periphery relations within the restructured education state that steers at a distance through outcome accountability measures. This move is linked to the transition from the Keynesian welfare state to the neo-liberal state (and austerity in some nations) that gives emphasis to outcomes rather than inputs, as with the Keynesian approach. Speaking more broadly about the rise of an audit culture related to state restructuring, Power (1997, p.44) sees the hollowing-out of state structures as generating 'a demand for audit and other forms of evaluation and inspection to fill the hole'. Test-based accountabilities are central to the culture

of performativity; in Lyotard's terms the input-output equations that in a cybernetic manner now keep educational systems operant, as (some) meta-narratives have passed away. These accountabilities are also central to choice discourses: public display and availability of school test data are seen as a component of the system of school choice with data transparency seen as central to the practices of choice. I note though, that not all parents and families are able to exercise choice in respect of schooling, further entrenching the advantages in schooling of the middle classes and better off, though we need to recognize that the middle classes have also become insecure in the new competitiveness that suffuses the contemporary structure of feeling.

Mention has been made of the new accountabilities linked to testing and standards as being a globalized education policy discourse. This is an important point, as we need to acknowledge the complementarity between international testing of schooling systems and national testing. There is a way in which they now sit in a symbiotic relationship with each other as part of what I referred to above as the fourth message system of schooling. In research I have done at the OECD, which conducts PISA every three years, a senior policy maker made the point that the two modes of testing are necessary, given improvements in national test performance without international comparison might very well mask decline in terms of international comparisons (Sellar & Lingard, 2013b). PISA, as well as the IEA's TIMSS and PIRLS, have become major policy levers within nations, often used as a form of 'externalization' (Schriewer, 1990) to justify national reforms rather than policy learning, a rationale articulated by the OECD for PISA testing. PISA 'shocks' of various kinds have reframed national policy developments in education. Think of the effect of Germany's poor performance on the initial PISA in 2000 (Grek, 2009). This led to a plethora of educational reforms; equally, we could think of the impact of Shanghai's stellar performance on the 2009 and 2012 PISA (Sellar & Lingard, 2013a). President Obama in the USA commented, in a state of the union address, that China outperforming the US on PISA to such an extent as it did, was this US generations' Sputnik moment. The then Australian Prime Minister, Julia Gillard, argued that Australia was in a danger of becoming the 'runt in the litter', set against the rise of China in the so-called 'Asian Century'. A consequence was the goal of Australia being back in the top five on PISA by 2015 being built in to legislation. The 2012 PISA results witnessed Australia moving further away from achieving this goal. We also know from the literature the distorting effects in education of such targets (e.g. Stobart, 2008).

The other point that needs to be made about the complementarity of international and national testing and related modes of accountability is that they are framed by similar ontologies and epistemologies. Indeed, they function through and are facilitated by global and national epistemic communities (Kallo, 2009) and networks of policy makers and new technicians that stretch across the globe (Ball & Junemann, 2012). We might speak here of epistemological governance functioning globally. These epistemic communities help make the multiple policy spaces associated with globalization legible for governing, some of which are topological rather than typographical in

character (see Allen & Cochrane, 2010). As Porter (1994) noted, what these epistemic communities do is help create ‘new entities, made impersonal and (in this sense) objective when widely scattered people are induced to count, measure, and calculate in the same way’ (p.390).

The enhanced policy role of national testing is linked *inter alia* to two broader phenomena, one related to the restructured state and its new *modus operandi*, the other to a broader socio-cultural change. The first is the rise of policy as numbers (Rose, 1999, Ozga & Lingard, 2007, Ozga, 2009, Lingard, 2011). Systems are now largely steered through numbers, statistics, data and the like. This has witnessed the rise of a new neo-positivism in policy production, in the policy work of the education state with implications for educational research (Lingard, 2013). This policy as numbers works at all levels of the new multiscale modes of governance associated with globalization, what we might from another perspective regard as the new spatialisations (and topologies) of governance inherent in globalization (Amin, 2002). The second phenomenon is what Lyotard (1984) described as the death of meta-narratives in the wake of the end of the Cold War. In Lyotard’s terms, this death has seen the resurrection of new performativity as the mover of the machine of state. Input-output equations, data, metrics, statistics and the like fill the black hole left by the passing of meta-narratives and keep the system functioning. There is a second stage of this culture of performativity currently emerging in the big data phenomena, whereby census rather than sample data of multiple kinds are being used to manage systems with a focus on predictability and taming both the future and chance.

Both phenomena are reflected in the habitus of senior policy makers in schooling systems at various levels of the nation, but also in international organisations such as the OECD. This similarity of habitus across scales is central to the epistemic communities that manage testing at all policy levels across the space of the globe. This habitus sees the globe as a commensurate space of measurement, a stance that elides, discursively at least, the cultural idiosyncrasies of nations. Strathern (2000) has argued that globalization can be conceived in one way as the capacity (habitus in my argument) of individuals to imagine and think in terms of a global space. Specifically, she refers to the global as ‘the infinitely recurring *possibility* of measurement – not the scales *but the capacity to imagine them*’ (Strathern, 2000, p.17). This habitus is now central to the logics of practice, in Bourdieu’s terms, of national education policy fields, but also to the emergent global education policy field (Lingard and Rawolle, 2011) and central to comparison as a mode of governance. This habitus is able to imagine the imbrications of local and other spaces and able to make various spaces and places legible for governing through a policy as numbers approach. Within topological constructions, the sense of what is ‘near and far’ is disrupted (Allen and Cochrane, 2010, p.1073), as the space of the globe comes under the purview, the gaze, of policy makers. This gaze in Brighenti’s (2007) frame might be seen as an ‘epistemology of seeing’ that defines ‘fields of visibility’ open to metrics, judgements and comparisons. This is a frame that Michael Simmonds and P. Taylor Webb use instructively in their

paper in this special number, linked to what they call ‘accountability synopticism’. Within this epistemology, these policy makers are able to see and imagine the globe and other policy spaces as commensurate and productive of comparative measures of performance of national schooling systems, where comparison is now central to governance.

Related, Maarten Simons (2013) suggests the disposition (*habitus*) of policy makers is now about looking around, comparing and learning from others, and positioning one’s own system and its performance globally; tests as global positioning devices. He describes this disposition as ‘globalist’ rather than ‘modernist’. Wiseman (2010, p.8) has observed that test-based accountabilities are linked to the concept of evidence-based policy (only ever evidence-informed in my view). He also comments insightfully on the extended space of data used now for policy making, which is linked I would aver to the *habitus* of senior policy makers, both national and international. Wiseman observes:

...what widely available international data on education has done is create an intellectual space where educational policy-making is not geographically or politically bounded but is instead bounded by the extent of the legitimated evidence used to support one decision or policy versus another. (Wiseman, 2010, p.8)

Simons (2013) has argued that the rise of evidence-based or-informed policy has not only drawn upon hard data, but also on examples. He refers to this as ‘soft evidence’, as opposed to the ‘hard evidence’ of policy as numbers. Soft evidence includes examples of good practice, narrative accounts and so on, but as with hard evidence, is central to contemporary governance, argues Simons, through the ‘power of the example’. This is governance through comparison of soft data (as with hard data) and also reflected in the *habitus* of policy makers and the logics of practice of contemporary policy making in education.

There is probably now a disconnection, almost an incommensurability, between the *habitus* and disposition of high level policy makers and those who work in schools, a disjunction between global spaces and local places. Stevenson and Wood document this in their paper. We might say there are different spatio-temporalities framing the work of each and reflected in their different *habitus*. In an almost Orwellian sense, those working in schools might be seen as variables in a large data set to be managed, manipulated and measured to enhance student performance, the dependent variable in this equation, as we witness the emergence of a Deleuzian ‘control society’ of perpetual evaluation, the ‘age of infinite examination’ in Foucault’s terms.

One thing that the availability of school performance data has done is open up a new and potentially lucrative market for what have been called ‘edu-businesses’ (Burch, 2009, Ball, 2012). Privatisations have been central to the work of the state in these neo-liberal times (Ball, 2007, Burch, 2009), with testing opening up markets for the likes

of Pearson and other edu-businesses, in terms of the actual production and analysis of tests, but also in respect of the sale of professional materials for teachers in terms of testing regimes and primers for parents. All of the OECD's PISA data and analyses are freely and publicly available, and this is a good thing, but this situation also opens upon the possibilities of such data being utilised by edu-businesses, as is indicated by Pearson's *The Learning Curve*, which uses international performance data to establish policy problems for systems. This problem construction by Pearson, a feature of all policy making, and the reason why deconstructing the policy problem needs to be the first step in critical policy analysis, potentially creates markets for Pearson, who are changing business focus from publications to the provision of educational services (see Hogan, Sellar & Lingard, 2013). Test-based accountabilities in schooling systems provide fertile grounds for privatizations and edu-businesses. In an interesting variation on this privatization theme, Simmonds and Webb show how a think-tank (the neo-liberal Vancouver-based Fraser Institute committed to markets and deregulation) and the media were central to the construction of a quasi-market and school choice in the Canadian province of British Columbia. They did so through publication of league tables of school performance, preempting developments by policy makers.

The next set of issues in respect of test-based accountabilities that I want to address relates to their multiple impacts in various parts of schooling systems. Either explicitly or implicitly, all of the papers in this special number deal with these important issues for educators. Politically and systemically, test-based data function as catalysts with effects across systems and into schools and classrooms (Lingard & Sellar, 2012). Teachers sit at the interplay, the fold if you like, between external policy and internal classroom pedagogies. Testing has systemic effects through concerns for reputational damage; often with resultant political pressures focused on targeted improvement on testing that we know has perverse and anti-educational effects (Stobart, 2008, Lingard & Sellar, 2013). These catalyst effects work in respect of both international and national testing. There is also a plethora of research demonstrating the perverse effects on teacher work, perhaps best encapsulated in Taubman's (2009) talk of 'teaching by numbers', but also in the encouragement of scripted rather than intellectually demanding pedagogies, what we might see as 'pedagogies of indifference' (Lingard, 2007). There is also a way in which high stakes testing 'responsibilizes' teachers, especially when testing is accompanied by 'value-added' approaches and statistically similar school measures, which seek to control for contextual and non-teacher effects. There are also impacts on curriculum, particularly reducing the width of curriculum in those schools and classrooms populated by the most disadvantaged students who tend to not do so well on standardized tests: a new form of curriculum disadvantage and hierarchising of the distribution of powerful curriculum knowledges and capitals reinforcing class-based reproduction. Greg Thompson's paper, while indicating that some teachers made positive comments on the National Assessment Program – Literacy and Numeracy (NAPLAN) concerned with the encouragement of a coordinated cross-school focus on literacy and numeracy, more collaboration and resource sharing amongst teachers, and allowed for better tracking and monitoring of student progress, most talked about these

negative impacts. Stevenson and Wood worryingly illustrate how the policy control of teachers' work through testing also, through its responsabilizing and individualizing of teachers' work, limits the potential collective resistance against these developments by Teacher Unions.

Au and Thomas in their papers show the significance of the history of testing in the US to what is happening in the current accountability era. Au documents how the development of tests was linked to a meritocratic ideology, but that at the same time testing legitimated class and race privileges. He demonstrates how the 'assumptive objectivity' of the tests reinforced inequalities around race, class, poverty and gender, as the 'norming' of the tests was based on existing social and economic hierarchies. Additionally, he notes the paradox that test-based reforms are central to the political agenda to close the performance gap across the racial divide, while neglecting structural inequality and historical oppression.

Within the dominant human capital framing of education policy - the economization of education policy – Thomas argues that the reproduction of capitalism rather than democratic imperatives now drive school reforms based around high stakes testing. Both Au and Thomas also show how Reagan's *A Nation at Risk* (1983) saw the growth of high stakes testing in the states of the US with impacts on teachers and classrooms. According to AU, by 2000 every state in the US except Iowa had state mandated tests. Policy developments in schooling under Bush and Obama have seen an enhanced federal government presence in respect of testing and accountability.

I would note that the report *A Nation at Risk* and these subsequent developments in testing were central contributing factors behind the US's push at the OECD for the development of what became PISA, first administered in 2000 and then subsequently every three years. The US saw international comparative performance measures of their schooling system as central to strengthening their political and economic position within the global economy. In a sense as well, such tests provided a comparative measure of the nation's human capital and thus a surrogate measure of the economy's likely future global competitiveness. Thomas and Au also show how *No Child Left Behind* and Obama's *Race to the Top* are 'logical' outcomes of the history of testing in the US. We also see a strengthened federal government presence and Governors agreeing to national standards in Maths and English. The testing reforms and their effects that Thompson outlines in his paper have also affected the functioning of educational federalism in Australia with the attempt to constitute a national system of schooling, when Constitutionally the states and territories are responsible for its delivery. Testing is central here as it is a technology of distance.

Both Au and Thomas importantly document the ways in which schooling systems driven by testing function to mask the impact of racism, poverty and social inequality on schools and student learning. Indeed, one important finding of OECD analysis of varying national performance on PISA is that quality and equity can and should work together. Furthermore, the nations that do best on PISA have low Gini Coefficients of

inequality, a reality silenced in most analyses and accounts of comparative performance on the test. Likewise, we know that student background contributes much to school performance, while contemporary test-based accountability cultures, as I argued above, responsabilize teachers, denying and masking contextual effects and growing structural inequality. This is a point very well made by both Au and Thomas.

IN/CONCLUSION

The papers in this special number then make a most productive contribution to debates about the genealogy and effects of test-based accountabilities. They strongly suggest the necessity of historicizing and contextualising these developments for enhancing understanding of what is going on. In a sense, each documents the effects of this new-old system of testing. At the same time, each contribution also shows quite clearly how the globalized educational policy discourse of test and standards driving up performance manifests in vernacular ways within nations and provinces. Politics are part of this mediation, as John O'Neill so clearly documents in his account of why New Zealand has standards-based accountability, rather than a test-based form. The papers raise issues of political resistance to such developments, including the increasing involvement of edu-businesses and 'business capital' in this hegemonic policy regime and related democratic deficit. A new imaginary is surely needed to frame possible new policy regimes in schooling and will involve debates over: what should count in schooling; how richer, more intelligent modes of educational accountability might be created that hold politicians (not only teachers) to account; what might horizontal relations of accountability between schools and their communities look like; who should control the field of judgement; and how structural inequality needs to be addressed in both broad social policy and in education. Within that new imaginary how should we conceive of a new form of teacher professionalism, based on trust, rather than distrust, which seems to underpin contemporary moves to test-based accountability? What ought schooling be today, what ought its goals be, in terms of the types of future societies we want, locally, nationally and globally? What is the role of testing in respect of such debates and in a more socially just and democratic schooling system? The papers in this special number make an important contribution to opening up that necessary political conversation.

REFERENCES

- Allen, J. & Cochrane, A. (2010). Assemblages of state power: Topological shifts in the organization of government and politics. *Antipode*, 42(5), 1071-1089.
- Amin, A. (2002). Spatialities of globalization. *Environment and Planning*, 34(3), 385-399.
- Ball, S.J. (2013). *The education debate*. (2nd ed.) Bristol: The Policy Press.
- Ball, S.J. (2007). *Education plc*. London: Routledge.
- Ball, S.J. (2012). *Global education Inc*. London: Routledge.

- Ball, S.J. & Junemann, A. (2012). *Networks, new governance and education*. Bristol: The Policy Press.
- Bernstein, B. (1971). On the classification and framing of educational knowledge. In M.F.D Young (Ed.), *Knowledge and control*. (pp. 245-270). London: Collier-Macmillan.
- Brighenti, A. (2007). Visibility – a category for the social sciences. *Current Sociology*, 55(3), 323-342.
- Burch, P. (2009). *Hidden markets: The new educational privatization*. New York: Routledge.
- Connell, R.W. (1985). *Teachers' work*. Sydney: Allen & Unwin.
- Grek, S. (2007). Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy*, 24(1), 23-37.
- Hogan, A., Sellar, S. & Lingard, B. (2013). The learning curve: Pearson, edu-business and new policy spaces in education. Unpublished paper presented at The University of Queensland, Brisbane.
- Kallo, J. (2009). *OECD education policy: A comparative and historical study focusing on the thematic reviews of tertiary education*. Jyvaskyla: Jyvaskyla University Press.
- Lingard, B. (2007). Pedagogies of indifference. *International Journal of Inclusive Education*, 11, 245-266.
- Lingard, B. (2011). Policy as numbers: Ac/counting for educational research. *The Australian Educational Researcher*, 38(4), 355-382.
- Lingard, B. (2013). The impact of research on education policy in an era of evidence-based policy. *Critical Studies in Education*, 54(2), 113-131.
- Lingard, B. & Rawolle, S. (2011). New scalar politics: implications for education policy. *Comparative Education*, 47(4), 489-502.
- Lingard, B., Martino, W. & Rezai-Rashti, G. (2013). Testing regimes, accountabilities and education policy: commensurate global and national developments. *Journal of Education Policy*, 28(5), 539-556.
- Lyotard, J. (1984). *The Postmodern Condition*. Minneapolis: The University of Minnesota Press.
- Ozga, J. (2009). Governing education through data in England: From regulation to self-evaluation. *Journal of Education Policy*, 24(2), 149-162.
- Ozga, J. & Lingard, B. (2007). Globalisation, education policy and politics. In B. Lingard & J. Ozga, (Eds.) *RoutledgeFalmer reader education policy and politics*. (pp. 65-82). London: Routledge.
- Nichols, S & Berliner, D. (2007). *Collateral damage: How high stakes testing corrupts America's school*. Cambridge, MA: Harvard University Press.
- Porter, T. (1994). Making things quantitative. *Science in Context*, 7(3), 389-407.
- Power, M. (1997). *The audit society: Rituals of verification*. Oxford: Oxford University Press.
- Rizvi, F. & Lingard, B. (2010). *Globalizing education policy*. London: Routledge.
- Rose, N. (1999). *Powers of Freedom*. Cambridge: Cambridge University Press.