

COULD DO EVEN BETTER? Making the most of international comparison as a tool of policy

Robin Alexander

Background

As I have indicated to Tim Oates and Stefano Pozzi, I am keen to discuss the Department's current quest to identify how high performing school systems (as judged by their PISA rankings) describe and frame the curriculum (a) nationally and (b) within their schools and classrooms, and the kind of evidence which is being selected to support this exercise.

In this matter my interest is not confined to primary education or steered solely by the preoccupations of the Cambridge Primary Review. Instead - and I have to mention this because in DfE if not elsewhere I tend to be identified with English primary education - it stems from my research and consultancy over many years in comparative, international and development education. This has included work in countries/jurisdictions/education systems which of late have ranked high in PISA (Finland, Hong Kong, Singapore, Australia); countries with good but uneven PISA profiles (United States, Norway, Denmark, Germany, UK); one whose ranking has dropped dramatically because economic, political and social challenges have severely weakened an education system which was previously strong (Russia); and others outside PISA which face the characteristic challenges associated with medium and low HDI status (India, Bangladesh). In most of these cases I have covered the full spectrum from policy to practice, and have engaged directly with pedagogy through classroom observation, video recording, interviews and so on. I also bring perspectives from my membership of the comparative education community and my recent presidency of the British Association for International and Comparative Education.

Essential knowledge and curriculum design: learning from high performing systems

Following the general election and publication of Tim Oates' paper *Could Do Better*¹ the government announced that it intended:

- to cut back the statutory national curriculum to 'essential knowledge in the key subjects' and ensure that it is 'based on concepts, principles, fundamental operations and key knowledge [which] can lead to learning processes which are more focused on deep learning, [that is,] fewer topics pursued to greater depth'.²
- to benchmark England's national curriculum against the curricula of PISA high performers such as Singapore, Hong Kong and Finland, especially in respect of the curriculum priorities, structure and content which are presumed to contribute significantly to their success.

Though, as I note again later, the CPR under my leadership has taken a firm line on the centrality of knowledge and disciplinary understanding to a properly-conceived curriculum³ - often against the

¹ Oates, T. (2010) *Could Do Better? Using international comparisons to refine the National Curriculum in England*, Cambridge: Cambridge Assessment.

² The first quote comes from the national curriculum review's published remit, the second from Tim's paper.

³ Alexander, R.J. (ed) (2010) *Children, their World, their Education: final report and recommendations of the Cambridge Primary Review*, Routledge, pp 237-278.

grain of professional sentiment in the primary sector - I have also raised, publicly and privately, questions about the intentions summarised above.⁴ For example, in the absence of a clear and carefully-formulated set of educational aims, by what criteria do we define 'essential knowledge'? Can what is deemed essential in one culture be assumed to be equally essential in a very different culture? Without denying the importance of high standards and international competitiveness, should benchmarking against PISA high performers be the main starting point for constructing a national curriculum? How do we balance this international preoccupation with attention to national and local needs, opportunities, histories and cultural realities in a society which exhibits serious symptoms of dislocation? Why do we not pay more attention, in formulating a curriculum which is supposedly relevant to national needs, to other international league tables, for example on poverty, equality and wellbeing?

I and the CPR team have also pressed the educational and evidential case for children to have a statutory entitlement to a broad curriculum in which what constitutes breadth is specified rather than left to chance, and in which breadth is firmly allied to the highest possible quality of teaching and professional expertise in all subjects, not just those in the current national curriculum core and/or which feature in international achievement surveys like PISA, TIMSS and PIRLS.⁵

Incidentally, our argument in this campaign - for that is what it seems to have turned into - is as much about standards as entitlement. Given the weight of the Ofsted evidence about the direct relationship between curriculum breadth and standards in literacy and numeracy, not to mention strong corroborating evidence from the recent US report on arts education⁶, I'm baffled that DfE appears not yet to have publicly turned its back on the now discredited line that concentrating on 'the basics' and not worrying overmuch about the rest of the curriculum will deliver the standards we all wish for.

However, these are not my main concerns here, though being unresolved they remain prominently on file. Nor, despite its title, is this a critique of Tim's paper, though he knows that I'm rather more taken with some aspects of it than others and that although his paper ends with an impressively long list of references, I believe there are some significant omissions of work - both individual published studies and major generic lines of research - which really needed to be included in a desk-based analysis of the kind that he was undertaking, especially given the political decisions that the paper has been used to justify. Such omissions could well risk skewing the inferences that are drawn about cause and effect in other systems' educational performance, the recommendations for change that are made, and ultimately the quality of education that our children receive - though of course this is the responsibility of governments and schools rather than academic commentators. However, take this paper's title as an addendum to Tim's rather than a critique.

My concern, about lack of sufficient range and balance in the literature on which policy-directed inferences and recommendations are based, applies much more strongly to some other studies of this

⁴ For example: in speeches/presentations to the NAHT, ATL and NUT annual conferences, the Westminster Education Forum and at many other events; articles in *The Guardian* ('Give all children the riches they deserve', 15 March 2011) and *TES* ('Evidence from the US tells us the arts belong at the heart of the curriculum - but it's our last chance to make this happen', 19 August 2011 - the articles' titles were devised by sub-editors, not myself); also, and in greater depth, Alexander, R.J. (2010) ' "World class schools" - noble aspiration or globalised hokum?', BAICE Presidential Address at the 10th UKFIET Conference on Education and Development, *Compare: a journal of comparative education*, 40:6, pp 801-817; Alexander, R.J. (2008) 'Pedagogy for a runaway world' in *Essays on Pedagogy*, Routledge, pp 121-153; and of course in the interim curriculum report and the final report of the Cambridge Primary Review.

⁵ See the CPR's evidence to the first phase of the national curriculum review, its final report, and correspondence with the Secretary of State and the Minister for Schools.

⁶ Ofsted (1997) *National Curriculum Assessment Results and the Wider Curriculum at Key Stage 2*; Ofsted (2002) *The Curriculum in Successful Primary Schools*; Ofsted (2008) *Improving Primary Teachers' Subject Knowledge Across the Curriculum: a summary of evidence from subject surveys (excluding English and mathematics)*; Ofsted (2009) *Twenty Outstanding Primary Schools*, Ofsted (2010) *The Annual Report of Her Majesty's Chief Inspector of Education, Children's Services and Skills, 2009-10*; President's Committee on the Arts and Humanities (2011), *Reinvesting in Arts Education: winning America's future through creative schools*, Washington DC: RMC Research Corporation.

kind - for example the Ofsted-commissioned *Worlds Apart?* review by Reynolds and Farrell⁷ which attempted a similar job to Tim's on the basis of all international achievement data up to that point and had a considerable impact on the development of the previous government's national literacy and numeracy strategies, and Barber's reports for McKinsey⁸. Note that this is not a matter of the *number* of sources cited in such publications but their range and representativeness in terms of methods used and conclusions reached, their quality as judged by peer review, and the validity of their proposed application to education in the UK or elsewhere.

International comparison and educational policy: a typology

A succinct but authoritative discussion of different kinds of international comparative study, and their uses in the context of policy, was published in 2003 by the National Research Council of the National Academies, the US equivalent of the Royal Society, the British Academy and the Royal Society of Engineering (though collectively the US National Academies have a wider reach in terms of the professional and academic communities involved). Through its Board on International Comparative Studies in Education (no longer extant, though its work continues), the NRC has both advised and provided a commentary on US participation in the international student achievement surveys, and on related matters.⁹

The 2003 NRC report¹⁰ differentiates three main types of comparative study (and here I quote partly from the report in question and partly from my published 2009 BAICE Presidential Lecture¹¹):

Type I studies typically include large-scale surveys that aim to compare educational outcomes at various levels ... Type II studies are designed to inform one or more particular ... education policies by studying specific topics relevant to those policies and their implementation in other countries. Type III studies are not designed to make direct comparisons ... in terms of specific policies or educational outcomes. Rather, they aim to further understanding of educational processes in different cultural and national contexts. (National Research Council 2003, 13)

Type I includes the large-scale international student achievement studies like TIMSS, PISA and PIRLS. They are typically quantitative and referenced by a limited number of indicators and measures of learning outcome and, latterly, of context, input and process.

Type II includes a more varied mix of empirical studies - quantitative, qualitative, descriptive, interpretive - and literature reviews, though all with a relatively narrow focus in response to specific policy concerns or questions. *Worlds Apart?*, the McKinsey reports and *Could do Better* would come into this category. In search of answers to policy questions they may look at one system in detail or they may compare several.

Type III includes the majority of work in the published corpus of comparative education. These studies focus on one or more countries and their education systems in a more holistic way. They can be large or small scale, quantitative or qualitative (they are always empirical, never wholly literature-based) but though some may have significant policy applications the imperatives of policy are not their initial impetus. Rather, their goal is the advancement of understanding of other countries, their education systems, schools and classrooms for its own sake, and for what this tells us about ourselves.

⁷ Reynolds, D., Farrell, S. (1996) *Worlds Apart? A Review of International Surveys of Educational Achievement Involving England*, London: OFSTED

⁸ Barber, M. and Mourshed, M. (2007) *How the World's Best-Performing School Systems Come Out on Top*, Dubai: McKinsey and Company; Mourshed, M., Chijioko, C. and Barber, M. (2010) *How the World's Most Improved School Systems Keep Getting Better*, Dubai: McKinsey and Company.

⁹ See for example the publications listed at http://www.nap.edu/openbook.php?record_id=9903&displayrelated=1.

¹⁰ National Research Council (2003) *Understanding Others, Educating Ourselves: getting more from international comparative studies in education* (C.Chabott and E.J.Eliott, editors), Washington DC: The National Academies Press <http://www.nap.edu/openbook.php?isbn=0309088550>

¹¹ 'World class schools: noble aspiration or globalised hokum?', cited in other footnotes.

Being - in that wantonly myopic phrase - 'not invented here' (i.e. not officially commissioned or sanctioned) they tend simply to be ignored in departmental desk searches of the literature. However, since a good few have been funded from the public purse (for example with ESRC grants) this is not only a waste of potentially useful evidence and insight but also perhaps a dereliction of the duty of accountability.

Confirming these latter claims, NRC has no doubt where the power and perceived policy relevance lie, for while the majority of published comparative education studies are Type III, Type I and II studies receive most of the funding, political patronage and/or publicity, and the funding difference per study can be truly vast. Type I studies in particular are a multi-million dollar business, Type II may be generously or modestly funded, depending on whether they entail fieldwork or merely a literature review. But what they lack in funding they make up for in patronage. Type III studies get what they can from funding bodies, for such studies are rarely supported by government, directly at least. Yet, the NRC report goes on:

Although they vastly outnumber Type I and Type II studies, Type III studies often do not come to the attention of policy makers or the public. This is a loss, since many are rich in narrative detail and paint a more engaging and provocative portrait of education in other countries than do the summary bar charts and graphs typical of many larger studies. Ethnographic and case studies, in particular, can explore cultural context in depth and, in turn, help elucidate the way education is organised and understood in different cultures. (National Research Council 2003, 23-4)

Note those words 'provocative' and 'understood'. Rigorously undertaken (and this condition is by no means always met), comparative educational research probes surface structures, does not take official information at face value, triangulates viewpoints, probes insider perspectives, seeks to understand as well as inform, and addresses those questions about motive, power and consequence which a system's guardians might prefer to leave unasked. Especially, it remains alive to the risks of ethnocentrism, naivety or plain ignorance which always attend the act of portraying another culture from the outside, and it respects the academic principle of *cumulation*: seeking out and building on research that has gone before (which means engaging with it, not merely peppering one's writing with references gleaned from bibliographies and journal keyword searches).

This, of course, is the ideal. Some comparative research - and far too much educational research generally - falls short, especially as to the last condition above.

Missing perspectives, missing the point?

Of course, there are other ways of classifying comparative studies, and compared with several established typologies in the mainstream comparative education literature¹² the NRC one somewhat oversimplifies. But the issues to which the NRC typology alerts us are important:

- Although, for example, PISA has begun to open up cultural questions through its (as yet limited) data on equity, and the TIMSS video studies have gained a partial purchase on pedagogy, Type I studies do not and cannot provide the kind of close-grained engagement with culture, schools and classrooms which, at best, Type III studies offer. Yet it is only from such engagement that we gain real understanding of educational cause and effect, of the thinking and history that underpin the admired practice, and the prospects for imitating or emulating it elsewhere.

¹² For example: Cowen, R. and Kazamias, A.M. (ed) (2009) *International Handbook of Comparative Education*, Dordrecht: Springer; Bray, M., Adamson, B., Mason, M. (ed) (2007) *Comparative Education Research: approaches and methods*, Hong Kong and Dordrecht: Springer, 2007; Crossley, MW, Broadfoot, P & Schweisfurth, M. 'Changing educational contexts, issues and identities', in M Crossley, P Broadfoot & M Schweisfurth (Eds.), *Changing Educational Contexts, Issues and Identities. 40 Years of Comparative Education*, (pp. 1-15), Routledge, 2007. Broadfoot, P. (1999) 'Not so much a context, more a way of life? Comparative education in the 1990s', in Alexander, R.J., Broadfoot, P. and Phillips, D. (ed) *Learning from Comparing: new directions in comparative educational research*, Oxford: Symposium Books, pp 21-32; Altbach, P.G. and Kelly, G.P. (ed) (1986) *New Approaches to Comparative Education*, Chicago: University of Chicago Press.

- Those Type II studies that are desk-based lack the vital asset, possessed by both Type I and Type III studies, of first-hand empirical data systematically and transparently presented, and analysed in accordance with methodological procedures that are publicly recognised or, if they are experimental, are at least open to scrutiny. Although protocols have been developed, for example for the EPPI reviews, to try to ensure that literature-based studies engage comprehensively, even-handedly and rigorously with the literature in question (and we followed such protocols in the 28 research surveys commissioned by the Cambridge Primary Review¹³), too many desk-based Type II studies are notable for a high degree of selectivity, arbitrariness and even bias in the literature on which they draw and in the way they handle it. That, I'm afraid, is part of their attraction in policy circles, because they prove a point or come up with a solution rather than demonstrate the complexity which a properly-conducted comparative study cannot help but reveal. This kind of study has a further and particularly serious disadvantage. It characterises and comments on systems, schools and classrooms elsewhere second or third hand, without having the direct experience of observing and engaging with those who live and work there from which a properly nuanced understanding comes.
- In seeking to explain the differential performance of education systems in student achievement surveys like TIMSS and PISA, or in seeking to generate curriculum or other policies which respond to such surveys validly and with a reasonable chance of success, it is therefore absolutely essential for DfE to engage with all three kinds of study - and to do so armed with appropriate methodological expertise and understanding of the challenges and imperatives of cross-national comparison, and with critical antennae fully extended. I trust that in the context of the national curriculum review this is what DfE will try to do.

Beyond the information given

Let us turn now to the information about other educational systems on which the DfE draws or might draw.

- At national level there are readily available sources of the information required in the publications and websites of national governments and – for the purposes of comparison – NFER's INCA¹⁴, the EU's Eurydice¹⁵ and international organisations like the OECD, the World Bank and the UN.
- Although such information appears to have the virtue of being straightforwardly factual, it has been shown to be of mixed reliability, because in some countries the official providers of the requested information are not above presenting their countries' educational policies in the most flattering light. For example, during the 1990s one Eurydice publication told the world that England's national curriculum provided 'a broad and flexible framework' that enabled primary schools to make their own decisions about what and how to teach,¹⁶ a statement of supposed fact with which neither the present government nor teachers working during the 1990s would agree. In any event, and crucially, such national information invariably deals with official expectations or requirements rather than actual practice and, as we all know, the curriculum as specified can be very different from the curriculum as enacted by teachers and experienced by children. Yet it is the latter that matters most. In any case, in some cases – I instance the Hong Kong curriculum – the official requirements can be so complex as to be virtually impossible to implement as formally specified.

¹³ Alexander, R.J., Doddington, C., Gray, J., Hargreaves, L., Kershner, R. (ed) (2010) *The Cambridge Primary Review Research Surveys*, Routledge.

¹⁴ International Review of Curriculum and Assessment Archive, supported by NFER and QCDA and run by NFER: http://www.inca.org.uk/about_us.html
Information on European educational systems:
http://eacea.ec.europa.eu/education/eurydice/index_en.php

¹⁶ Eurydice (1994), *Pre-School and Primary Education in the European Union*, Brussels: Eurydice, p 77.

- The problem is compounded in some Type II studies by the way that official or officially-endorsed information and claims constitute their core database. From such material - high on political rhetoric, low on classroom practice, devoid of cultural context, cleansed of problematic realities - they then construct edifices for school improvement and systemic reform. Disarmingly straightforward and carefully pitched to appeal to political instincts, these studies are eagerly taken up in the corridors of power.
- Aside from what can be inferred from, for example, the TIMSS video studies of maths teaching in Japan and Germany, information on how national curriculum requirements are translated into school and classroom planning and practice is *not* generally available in the Type 1 surveys. It is provided in some Type 2 studies, but the richest source by far is Type 3. But Type 3 studies, as the NRC report points out, are largely ignored by governments. They also get short shrift in the Type 2 desk-based reviews that come to governments' attention, partly because governments of their nature are more interested in systems, partly because in many systems decisions about curriculum implementation and pedagogy are left to schools and teachers, and partly perhaps because the reviewers may be unaware of the literature's full scope.
- There is also a tendency in Type 2 desk reviews to be overly deferential towards official sources, generating the risk that policy avoids proper critique and simply validates and replicates itself. This is a variant on that common and ultimately fatal political ailment 'Not invented here.'
- The failure to give due attention to Type 3 studies, it has to be acknowledged, is partly the fault of the academic community itself. Such studies may lie buried in low-circulation and overpriced academic journals, and too few education academics recognise the importance of communicating their work in accessible form to policymakers and practitioners. Many, indeed, prefer to remain within their institutional cocoon and are fearful of sticking their heads above the parapet. Having said all that, Type 3 studies - as the NRC report reminds us - are now the richest source of insight into the way that in other countries the curriculum is realised and enacted in school and classroom.

Methodological issues and imperatives

So much for the information available to the Department. The information imbalance I have noted, surely, can now be rectified. What of the methodology of the exercise? In seeking to benchmark against or extrapolate from PISA high performers like Singapore and Hong Kong we should be aware of the following.

- PISA assesses the attainment of 15 year olds in aspects of reading, mathematics and science. Its spectrum of 'key competencies', though clearly essential, is actually quite limited. Unlike some who use it as a badge of national pre-eminence or a stick with which to beat teachers, PISA itself more realistically and modestly acknowledges that its tests cover just 'some of the knowledge and skills that are essential for full participation in society'.¹⁷ PISA is right: its surveys are not a sufficient basis for describing an education system *as a whole* as 'high performing'.
- In fact, the wider discourse of educational quality and standards relies overmuch on proxies. The CPR has objected that the KS2 SATs cannot be treated as proxies for the entirety of children's primary education or for standards of attainment across the board, and that the notion of 'standards' should cover children's progress and attainment in all aspects of the curriculum, not just in what is formally tested. To treat the SATs in this way both diminishes the quality of the wider curriculum and misleads parents.¹⁸ Similar distortion may follow from over-reliance on PISA and other international indicators of educational performance, notably those developed by OECD for its *Education at a Glance* series.¹⁹ PISA is a valuable tool, but let's keep it in perspective and not infer more from it than is legitimate.

¹⁷ http://www.pisa.oecd.org/pages/0,3417,en_32252351_32235918_1_1_1_1_1,00.html

¹⁸ Alexander (ed) (2010) *op cit*, chapter 23.

¹⁹ This problem, at international level, is examined in Alexander R.J. (2001), *Culture and Pedagogy: international comparisons in primary education*, Blackwell, chapter 1; Alexander, R.J. (2008) *Essays on*

- A curriculum is, or ought to be, a reflection of national culture, values and aspirations. Transposing a curriculum structure from one culture to another is therefore a risky and questionable business. Of course, it can be argued – though (NB) only up to a point – that subjects like maths and science are more universal and culture-neutral than, say, history or mother-tongue language and literature, but the exercise remains risky. And, actually, maths and science are far from value-free. Why is one culture obsessed with basic computation while another is no less interested in problem-solving? Why does one national curriculum focus exclusively on fundamental scientific propositions and processes in its account of ‘essential knowledge’ while another also engages students in the application of science to the challenges of human wellbeing and ecological sustainability? Choices in a science curriculum can actually be pretty political.
- Curriculum only acquires real meaning for students once it is enacted through pedagogy. But pedagogy is deeply redolent of national and local values, habits and histories, and indeed of national social and political relations. Put bluntly but without naming the system(s) in question, is it right - let alone viable - to import to England the curriculum and/or pedagogy of a one-party state in which dissent isn’t tolerated, merely because that state’s 15 year-olds outperform ours in PISA’s tests of aspects of reading, maths and science? (It is at least possible that authoritarian political structures breed an authoritarian pedagogy that tells students what to think rather than teaches them how to think for themselves).
- In any event, what of the Secretary of State’s assurance that pedagogy is once again a matter for teachers, not governments? How does this sit with the exercise of emulating the way PISA high performers structure the curriculum for teaching?
- Next, we return to what some call ‘cherry picking’, a tendency that publicly we deplore but then continue to pursue. This has a statistical variant: false correlation. Simply because x is a common feature of high-performing education systems a, b, c, d and e, that doesn’t demonstrate a cause-effect relationship between feature and performance. And if x is also a common feature of low-performing systems g, h, i, j and k, then the claimed relationship is clearly inadmissible. In 1996, the Ofsted *Worlds Apart* study found that in its chosen high performing systems (at that stage as judged by TIMSS results) whole class teaching was the main teaching method used. It therefore concluded that whole class teaching was one of the keys to delivering high standards. Hence its commendation in the NLS and NNS, prefaced by ‘interactive’ which is mere tautology since whole class teaching is by its nature interactive. However, whole class teaching is the international pedagogical default, as prevalent in low-performing systems as high, so actually there is no correlation. In any case, ‘whole class teaching’ covers a multitude of sins or virtues and what is really needed is close-grained and fully-contextualised study of the entire pedagogy of teachers working in the countries in question.²⁰ But this is not what Type 1 and Type 2 studies do; and Type 3 studies, as NRC records, don’t feature in policy deliberations.
- Further, if Type 2 studies of the kind that I understand the Department is interested in look *only* at high performing systems then they will deny themselves access to the controls which enable them to sift the genuine factor from the spurious one. It is essential to look at other countries too. Casting the net more widely and less exclusively may reveal, as exemplified, that characteristics which are claimed to drive up standards in high-performing systems are actually as prevalent elsewhere. Whole class teaching is one near-universal; an overriding emphasis on literacy and numeracy, as the Benavot studies have consistently shown, is another; so is a curriculum concentrating on what is deemed to be ‘essential knowledge in key subjects’ for that is the basis of

Pedagogy, Routledge, chapter 2; and, in greater detail, in a publication commissioned by DfID in the context of education reforms in India and that country’s prospects of meeting the UN Millennium Development Goal of universal primary education, in Alexander, R.J. (2008) *Education for All, the Quality Imperative and the Problem of Pedagogy*, DfID/IoE: CREATE.

²⁰ The 1996 Ofsted *Worlds Apart* study is subjected to a detailed critique in Alexander, R.J. (1996) *Other Primary Schools and Ours: hazards of international comparison*, Warwick, CREPE, updated in Alexander, R.J. (2008) ‘Pedagogy goes East’, *Essays on Pedagogy*, Routledge, pp 9-42.

most of the world's national curriculum specifications.²¹ In short, far from being unique to high-performing systems, the current curriculum holy grail of a narrow range of subjects dominated by literacy, numeracy and science is in fact a curriculum commonplace. The issue, of course, is what schools and teachers *do* with such specifications, which is all about pedagogy, Type 3 and the dilemma of government intervention.

- Finally, once we move beyond from the restricted range of system and school variables deployed in some of the Type 2 studies, most notably the McKinsey reports, we encounter wider social, cultural, demographic and economic conditions which undoubtedly influence the educational performance of a country's students. In a recent critique of such studies I have shown that while concluding that teachers and teaching make a considerable difference (true but utterly banal and certainly not worth McKinsey-level expenditure and hype), extra-educational factors like country size, per capita GDP, demographic homogeneity and relative equality correlate no less convincingly with PISA performance.²²

A century ago, Sir Michael Sadler – chair of a far-sighted commission on education in India, administrator, arts patron, vice-chancellor and comparativist – famously warned:

In studying foreign systems of education we should not forget that the things outside the schools matter even more than the things inside the schools, and govern and interpret the things inside ... The practical value of studying in a right spirit and with scholarly accuracy the working of foreign systems of education is that it will result in our being better fitted to study and understand our own ... No other nation, by imitating a little bit of German organisation, can thus hope to achieve a true reproduction of the spirit of German institutions ... All good and true education is an expression of national life and character.²³

It was Germany then: now it's Finland, Hong Kong and Singapore. In 2011, Sadler's proposition can be tested using data of a range and richness - about 'the things' both inside and outside the school - that he could not have anticipated in 1902. Provided, that is, that we are prepared to open our minds to what is available.

Conclusion

So, if the Department really wishes to make best use of evidence from international comparisons to guide its curriculum reforms, it might think about doing the following:

- Look across the full methodological spectrum of comparative studies, not just at Types 1 and 2.
- Give particular attention to the insights on *practice*, or curriculum as enacted and experienced - as opposed to policy, aspiration or official rhetoric - which is available from the best Type 3 studies.
- Treat all official 'information' about other education systems as claim to be verified rather than fact.
- Attend to other measures of national educational success and student wellbeing besides PISA, TIMSS and PIRLS; and without diminishing concern for literacy, numeracy and science use these

²¹ Benavot, A., Cha, Y-K., Kames, D., Meyer, J.W., Wong, S-Y (1991), 'Knowledge for the masses: world models and national curricula, 1920-1986', *American Sociological Review*, 56; Benavot, A. (2008) 'The organisation of school knowledge: official curricula in global perspective', in Resnick, J. (ed) *The Production of Knowledge in the Global Era*, Rotterdam and Taipei: Sense Publishers.

²² See the section 'Small, rich or equal?' in Alexander, R.J. (2010) ' "World class schools" - noble aspiration or globalised hokum?', BAICE Presidential Address at the 10th UKFIET Conference on Education and Development, *Compare: a journal of comparative education*, 40:6, pp 801-817. As an instance of what else might be factored in, see the discussion, based on extensive epidemiological evidence, on the relationship between equality and educational standards in Wilkinson, R. and Pickett, K. (2009) *The Spirit Level: why equality is better for everyone*, London: Allen Lane. I've already commented on the more extreme cases of Russia and India, and the dramatic decline of the former in the international achievement surveys.

²³ Sadler, M. [1900] 'How can we learn anything of practical value from the study of foreign systems of education?' In J.H.Higginson (ed) *Selections from Michael Sadler: studies in world citizenship*. Liverpool: DeJall and Meyorre.

other measures to frame a curriculum which responds more comprehensively to individual and national needs.

- Note, from both Type 1 and Type 3 studies, that ‘numeracy’ and ‘literacy’ are not fixed or universal entities and that, for example, some high performing systems take a much broader view of the latter – balancing and integrating it with oracy – than we do in Britain.
- Draw directly on the expertise and work of the UK’s comparative education community which, notwithstanding the government’s laudable emphasis on learning from other systems, has so far played little or no part in the national curriculum review.

Having said all this, I actually agree – on the basis of my own comparative research and my reading of the research of others – that curriculum focus and structure, and a strong epistemic framing to classroom practice, are essential ingredients of effective pedagogy. This was the line we took in the CPR final report. Had we ignored the research evidence and instead heeded only our primary teacher witnesses, we would have ended up recommending the kind of soft focus, epistemologically flabby ‘creative’ or ‘skills’ curriculum that is familiar to students of post-war educational history and is once again an alarmingly real prospect in some primary schools if the national curriculum requirements are restricted to the core subjects and it is left to schools to decide what else to do. Be warned.

So in essence I and NC review colleagues are on the same wavelength in terms of the desired epistemic character of the curriculum, and the issues on which we have yet to reach agreement concern (i) the scope of the statutory curriculum and (ii) what counts as evidence, what it is legitimate to infer, and what it is helpful to emulate, in studying education systems other than our own. I hope that this paper, taken with Tim’s, provides a basis for moving forward in an evidentially more secure way.

Note. By way of illustration of what a Type 3 study can offer, and on the particular issue of *structure* in curriculum and pedagogy, an appendix containing two extracts from *Culture and Pedagogy* is available on request.