

GENERAL COMMENTARY

Tim Oates

This book in general, but Tattersall's and Baird's chapters in particular, engages with vital issues of the way in which avowed concerns regarding comparability are located in economic, political, social and scientific development. Their analysis usefully draws our attention to the fact that there are different forms of comparability – at the very least between-board comparability, between-year and between-subject – and this variation combines with different conceptions of 'standard' as well as different conceptions of the purposes of assessment, of public examinations and of education. The potential and real complexity which emerges from this requires us to explain what it is that gives rise to preoccupations with certain forms of comparability at certain times. Shifting focus is demonstrated clearly through the last three decades, which have seen Carol Fitz-Gibbon's work on the possible movement of achievement standards in science A levels relative to other subjects (Fitz-Gibbon, 1991); concerns regarding the extent to which different A levels prepare learners for university-level education (for example, Beckett, 2006); the imperative for 'taking forward the standard' during times of system transition (exemplified in 2002 by the problems experienced with the wholesale Curriculum 2000 A level reforms); and the introduction of performance tables as a public policy tool for school accountability.

The latter is a perfect illustration of how shifts in 'comparability concerns' can be driven by things other than concerns within the assessment community to improve the technical performance of instruments and systems. The 'league tables' have at their heart embedded assumptions about the equal value of different qualifications in counting towards a school's overall performance table score. Classes of qualification tend to attract an identical score (all GCSEs get the same number of points); scores in different years are considered comparable and vital to judging a school's trajectory over time and its changing position relative to other schools. Between-subject comparability is presupposed in the model of performance tables which we have today, yet such comparability remains a contested and problematic issue. What's more, the performance tables increasingly have been expanded to include a very wide range of diverse qualifications – with contrasting origins, purposes, design models, assessment and grading systems. This has given rise to some significant anomalies and accompanying controversies about 'easier' routes (by schools) to high performance table scores (Revell, 2001) – with all the implications of 'falsely elevated position'. The peculiar, embedded assumptions regarding comparability within performance tables matter, since the performance tables are such a prominent mechanism, and carry real purchase on schools' reputations. For some, comparability may seem to be a desirable technical objective in its own right; others see it as a prerequisite of assessments which provide the foundation for much broader policy instruments.

The specific and peculiar factors which drive societal preoccupation with certain forms of comparability are capable of being examined and subjected to critique. This in turn gives rise to a key question – who should do this analysis and what approaches are most appropriate for it to be done with precision? Kathleen Tattersall's chapter gives one of the most comprehensive overviews of the development of public examinations which has been seen in recent years. Jo-Anne Baird's chapter highlights the issues of comparability which were at the heart of policy management of the complex system transition during the 2000 reforms of A levels. But this does give rise to an issue of method. Many of the key debates and decisions regarding comparability are embedded in 'grey' policy papers and in undocumented exchanges between policy-makers, officials and ministers. Baird connects conceptions of comparability with some of the seismic events in qualifications awarding in the UK, using these to illustrate the sociologically and politically embedded nature of these conceptions. Incisive analyses of these events are vital elements in the explanation not only of the implications of using different concepts of comparability, but also to the way in which the conceptions are themselves constituted. Who said what and who did what, and when, become important to scientific understanding of the conceptions of comparability. This invokes a requirement for formal historical, social and political analysis – with all the attendant requirements for sound investigative method. Currently, it is largely left to assessment specialists to produce accounts of events in their world. In one sense, they are the right people; they understand measurement theory and are immersed in the operation of systems. But things in the world of assessment are connected to, and determined by, social, economic and other drivers. To strengthen understanding and critique, the assessment community needs to encourage historians and sociologists to join in with the task of understanding movements and developments in assessment in general, and comparability in particular.

In trying recently to encourage leading sociologists and economists to present a challenging critique of assessment, I became aware that they felt an acute lack of confidence in being able to talk with any authority about assessment. There is no such reluctance on their part to engage with issues such as access to education, inequalities, etc. But assessment appears to be too technical, too specialist, for them to bring their own science to bear on it. This only leads to a diminution of our understanding of the historical, social and political location of assessment. Many have pointed to the complexity of adequate explanation in education and training, highlighting the need for multidisciplinary analysis in order to generate adequate understanding – for example, of incentives for participation (such as the ESRC Centre on Skills, Knowledge and Organisational Performance – SKOPE)¹, or of patterns of inequality (such as the ESRC Centre for Longitudinal Studies)². This book is more than 'a start' on comparability, but elements of it suggest that a wider community of analysts needs to be created to fully reveal the mechanisms at work. It could be argued that this deeper understanding and transparency is essential for creation of effective public policy.

Of immense import, Baird's and Newton's analysis suggests that adoption of an operational definition of comparability is not determined by the selection of an

approach which rises to the surface as a result of superior technical characteristics. Rather, approaches are adopted as a result of compromise, or deliberate choice, over purpose. So getting the historical detail right about who said and did what, when, is indeed crucial. Who decides on purpose, who asserts the choice of approach, who determines the mechanisms by which the results are asserted on learners and society, all count. Public accountability is not simple, it is a complex, fraught and mediated business.

The power relations around selection of purpose are thus crucial. Baird helpfully asks questions such as whose views have predominated in current arrangements, and why? What are the dominant social views of which form(s) of comparability are of most significance for society and the economy? These are empirical questions, almost certainly best explored by specialists in social systems. Beyond these questions of fact, whose voice(s) should determine revised arrangements? It should be remembered that the education and training system in England is diverse in form, function and content. GCSE and GCE as classes of qualification may dominate public discourse, but among these categories there is a wide variety of modes of assessment, and different balances of different modes. Beyond and within schooling, vocational and occupational qualifications possess very different models of assessment. The conceptions and apparatus of comparability vary here to a greater extent than in GCSE and GCE – to the extent that comparability is a grossly under-developed area in some compartments of the system. These system discrepancies are illuminating – why is comparability less of a (social and political) issue in some parts of the education and training system, and is this lack of prominence tenable; does it constitute an implicit neglect of the interests of certain learners/candidates? Or does this lack of concern hint at an over-developed sense of concern for certain types of comparability in assessment within mainstream education?

Endnotes

- 1 SKOPE homepage: <http://www.skope.ox.ac.uk/> (accessed 13 October 2007).
- 2 CLS homepage: <http://www.cls.ioe.ac.uk/> (accessed 13 October July 2007).

References

- Beckett, F. (2006, August 29). 'Website says no'. *The Guardian*.
- Fitz-Gibbon, C.T. (1991). A levels: Corrective comparisons. *Managing Schools Today*, 2, 44–45.
- Revell, P. (2001, August 21). Couldn't do better. *The Guardian*.