

Editorial

Tom Bramley *Assistant Director, Research Division, Assessment Research & Development*

In 2007 QCA published the book 'Techniques for monitoring the comparability of examination standards', the purpose of which was not only to provide a review of the comparability research carried out since a previous review in 1985, but also to describe and evaluate in more depth the different methodological approaches used to investigate comparability. Since that publication, the profile of comparability research has remained high with the formation of the qualifications and examinations regulator Ofqual, which in 2010 began a programme to compare the demand of qualifications and assessments internationally, focusing initially at pre-university level.

Given the central importance of comparability and standards to all aspects of Cambridge Assessment's work across its three business streams (OCR, CIE and ESOL), in 2008 Cambridge Assessment set up a Comparability Programme with three full time dedicated members of staff and an associated governance group in order to contribute to, and maintain an overview of, the wide range of comparability work carried out across the Group.

In this Special Issue of *Research Matters* we present some of Cambridge Assessment's recent thinking about comparability. In the opening article, Gill Elliott, leader of the Comparability Programme, gives an historical overview of comparability concerns showing how they have been expressed in different political and educational contexts in England over the last 100 years.

It has become increasingly clear that comparability research is bedevilled by a lack of clarity over the meaning of its most basic concepts and inconsistent use of terminology. The second article, also by Gill Elliott, identifies and defines some widely used terms and shows how different methods of investigating comparability can be related to different definitions.

A topic of perennial interest is the inexorable rise over the last 25 years in the percentage of students passing, or achieving A grades in, A level examinations. In the third article, Paul Newton, Director of the Cambridge Assessment Network, tries to find evidence to support the popular (mis)-conception that A levels used to be norm-referenced but became criterion-referenced, and that this change was responsible for the rising pass rate.

Another topic of recurring interest is whether, within a qualification type (e.g. GCSE or A level), subjects differ in difficulty. It always seems to have been easier to calculate indices of relative subject difficulty than to explain exactly what they mean. A recent approach has been to use the techniques of Item Response Theory, treating different exam subjects like different questions (items) on a test. In the fourth article I discuss whether this analogy works.

It is an unavoidable fact of comparability research that often there is a need to compare things that are in many ways very different, such as vocational and academic qualifications. A sensible basis for comparison needs to be found, and in the fifth article Jackie Greatorex discusses one such basis – 'returns to qualifications' – that has so far been relatively rarely used by researchers in awarding bodies. The appendices to her article (pp.39–41) include two glossaries, one of qualification

types and one of assessment terms, which readers unfamiliar with the acronyms and jargon of assessment in England may find useful.

In the world of language testing, comparability is perhaps best conceived as an aspect of validity – that is, comparability of inferences that are justified about the communicative competence of individuals with certificates from different language testing organisations. In order to bring some coherence to a potentially conflicting area, the Common European Framework of Reference for Languages (CEFR), published in 2001, was devised, with consequent political and commercial pressure for language testing organisations to map their own tests to the proficiency levels in this framework. In the sixth article Neil Jones, Assistant Director of Research & Validation for Cambridge ESOL, discusses some of the conceptual issues involved in linking tests to the CEFR.

Frequent change has been a constant feature of school examinations in England for many years. The most recent innovation, which appears at the time of writing likely to be short-lived, is the 'unitisation' of GCSE examinations. Whereas GCSEs were formerly taken 'linearly' by students aged 16 at the end of a two year course, now the different 'units' can be taken at various stages throughout the course. This naturally presented a great challenge to the exam boards to ensure that the outcomes on the first large-scale award of the new unitised examinations in June 2011 were in some sense comparable to those on the old linear ones. In the seventh article, Mike Forster, Head of Research & Technical Standards at OCR, describes some of the issues that arose, and research undertaken by OCR in order to develop guidelines for grading procedures in 2011 that would be capable of achieving comparability.

I suspect that few researchers in comparability would deny that the audience for their academic articles is relatively small, comprising a fairly closed circle of individuals writing mostly for each other's benefit. Many comparability stories that make the headlines, on the other hand, come from outside academia. The final article, by Nicky Rushton, Matt Haigh and Gill Elliott, takes an interesting step away from the academic literature on comparability to discuss how comparability issues are presented in the media, and to evaluate the contribution that programmes like "That'll Teach 'em" can make to our understanding of comparability and standards.

It is our hope that this Special Issue will be both thought-provoking and informative, and a useful point of reference for anyone interested in the complex and challenging issues of comparability of examinations and qualifications.

Acknowledgements

I would like to thank the following people* for reviewing the articles in this Special Issue: Beth Black, Robert Coe (CEM Centre, Durham University), Vicki Crisp, Mark Dowling, Gill Elliott, Mike Forster, Jackie Greatorex, Sylvia Green, Matt Haigh, Neil Jones, Paul Newton, Tim Oates, Helen Patrick (independent consultant) Nick Raikes, Stuart Shaw, Irenka Suto, and Carmen Vidal Rodeiro.

Tom Bramley

Cambridge Assessment, October 2011

*Cambridge Assessment unless otherwise stated.