



- 1 **Foreword** : Tim Oates, CBE
- 1 **Editorial** : Sylvia Green
- 2 **Revisiting the topics taught as part of an OCR History qualification** : Karen Dunn, Ellie Darlington and Tom Benton
- 9 **Accounting for students' mathematical preparedness for Finance and Business degrees** : Ellie Darlington and Jessica Bowyer
- 17 **Collaboration in the 21st century : Implications for assessment** : Simon Child and Stuart Shaw
- 23 **The effect of subject choice on the apparent relative difficulty of different subjects** : Tom Bramley
- 27 **On the impact of aligning the difficulty of GCSE subjects on aggregated measures of pupil and school performance** : Tom Benton
- 30 **Statistical moderation of school-based assessment in GCSEs** : Joanna Williamson
- 37 **Good – better – best? Identifying highest performing jurisdictions** : Gill Elliott
- 39 **Research News** : Karen Barden
- 41 **Statistics Reports** : The Research Division
- 42 **Introducing Data Bytes** : James Keirstead, Tom Sutch and Nicole Klir

If you would like to comment on any of the articles in this issue, please contact Sylvia Green – Director, Research Division. Email: researchprogrammes@cambridgeassessment.org.uk

The full issue of *Research Matters 22* and all previous issues are available from our website: www.cambridgeassessment.org.uk/our-research

Research Matters : 22

A CAMBRIDGE ASSESSMENT PUBLICATION

Foreword

The research which underpins and consolidates the quality and operation of our assessments remains fundamental to our programme of work, but we continue to combine this with contextual analysis regarding learning, participation and policy. These matters have high prominence in this edition of *Research Matters*, with two articles of particular significance, despite relatively modest titles. Elliott's article reporting her incisive work on identifying the highest performing systems brings some much-needed sophistication to the contemporary research scene associated with high-performing jurisdictions. Following the Programme for International Student Assessment (PISA) 2000 the 'done thing' was to buy a ticket to Finland. It seemed obvious. In fact, doing this was steeped in naiveté, as my own *Finnish Fairy Stories* (Oates, 2015) and Salhgren's excellent *Real Finnish Lessons* (Centre for Policy Studies, 2015) now have made abundantly clear. There is much to learn from Finland, if the right questions are asked, and the right 'guided enquiry' undertaken. Elliott's scrutiny of indicators does not undermine sound transnational comparative analysis; it supports sophisticated understanding of the assets and deficits of specific sets of national education arrangements. It does not ignore complexity, but indicates how an appreciation of the complexity of system performance can further the insights gained from comparative analysis. Meanwhile, back in our domestic system, the issue of 'all subjects being at the same level of difficulty' has emerged again in policy discussions, not least because of equivalence assumed in, and driven by, accountability arrangements. Work by Professor Coe, Director of the Centre for Evaluation and Monitoring, has examined the relative movement of standards in different subjects and we, at Cambridge Assessment, have argued that the pursuit of 'same difficulty' is not quite the obvious good which it might seem. Benton's analytic piece explores one dimension of this pursuit and the implications which it carries. As with all 'obvious' policy moves, this is one which needs thorough examination before rushing to immediate action.

Tim Oates *Group Director, Assessment Research and Development*

Editorial

The first two articles in this issue report on research that investigates subject content and skills in very different contexts. The research from Dunn, Darlington and Benton follows from the work of Child, Darlington and Gill (2014, 2015). This more recent research focuses on the topics that schools intended to teach as part of their A level History qualification (first teaching from September 2015). It provides interesting insights into the potential impact of reform of the A level History specification. Also in the context of reform, Darlington and Bowyer consider the impact of changes to A level Mathematics and A level Further Mathematics as a result of the reform programme. They discuss the changes to optionality and preparedness of A level students who proceed to study Business Studies at university.

Child and Shaw discuss process and outcomes in the context of a key 21st century skill, namely, collaboration. They recognise the importance of construct definition and the challenges related to validity, reliability, comparability, and delivery in assessment. This article informs debate on issues of construct definition, and task design, as well as the challenge of group assessment.

The next two articles consider the matter of subject difficulty from different perspectives. Bramley reports on his research into the thorny issue of whether and how it could be possible to control for inter-subject comparability. He used simulated data to investigate the validity of one statistical method and highlights the problem of trying to adjust for differences in difficulty at subject level. In his article, Benton argues that differences in subject difficulty do not cause problems for school accountability, or for summarising the achievement of students at GCSE. He used data from the National Pupil Database to support his conclusion and expresses concerns about the rationale for attempting to make different subjects 'equally difficult'.

The final two articles have an international flavour. Williamson explores the challenging area of statistical moderation of school-based assessment. She outlines methods of statistical moderation that are used in jurisdictions around the world and applies them to GCSE results data. Her work illustrates that further research is needed in order to reassure stakeholders before such changes to moderation processes could be considered. In the final article, Elliott identifies a challenge for those studying international comparability. There are many different comparisons that can be used to identify high-performing jurisdictions and this makes it increasingly difficult to identify a smaller number of them since their numbers grow with the number of comparative ranking exercises that are carried out. Elliott describes a definition that has been used in research at Cambridge Assessment to identify the highest performing jurisdictions. Although she identifies limitations to her approach, it provides an interesting and pragmatic definition to those wanting to identify a small number of the highest performers.

Sylvia Green *Director, Research Division*