



- 1 **Foreword** : Tim Oates
- 1 **Editorial** : Sylvia Green
- 2 **Extended essay marking on screen: Does marking mode influence marking outcomes and processes?** : Hannah Shiell, Martin Johnson, Rebecca Hopkin, Rita Nádas and John Bell
- 7 **The effects of GCSE modularisation: a comparison between modular and linear examinations in secondary education** : Carmen L. Vidal Rodeiro and Rita Nádas
- 14 **Tracing the evolution of validity in educational measurement: past issues and contemporary challenges** : Stuart Shaw and Victoria Crisp
- 20 **Does doing Critical Thinking AS level confer any advantage for candidates in their performance on other A levels?** : Beth Black and Tim Gill
- 25 **Comparing the demand of syllabus content in the context of vocational qualifications: literature, theory and method** : Nadežda Novaković and Jackie Greatorex
- 32 **Rank ordering and paired comparisons – the way Cambridge Assessment is using them in operational and experimental work** : Tom Bramley and Tim Oates
- 35 **A better approach to regulating qualification standards** : Bene't Steinberg and Sarah Hyder
- 38 **Statistical Reports** : The Statistics Team
- 39 **Research News**

If you would like to comment on any of the articles in this issue, please contact Sylvia Green.

Email: [researchprogrammes@cambridgeassessment.org.uk](mailto:researchprogrammes@cambridgeassessment.org.uk)

The full issue and previous issues are available on our website:  
[www.cambridgeassessment.org.uk/ca/Our\\_Services/Research](http://www.cambridgeassessment.org.uk/ca/Our_Services/Research)

# Research Matters : 11

A CAMBRIDGE ASSESSMENT PUBLICATION

## Foreword

*Research Matters* again establishes the importance of a highly elaborated research enterprise around assessment and qualifications. Shaw and Crisp's work on approaches to validity research is consistent with the commitments of the Cambridge Approach regarding the necessity for validation research; Shiell *et al.* drill down into issues of the impact of marking mode on reliability (again a commitment in the Cambridge Approach); and Novaković and Greatorex revisit the issue of demand – this time in the context of vocational qualifications. The work indicates the scale and extent of awarding bodies' responsibility for understanding the operation and impact of the assessments and qualifications which they develop and administer. Such work is not demanded in the quality requirements enshrined in English regulatory arrangements, and shows how meeting strict moral and technical standards extends well beyond the obligations contained in regulation. It is on this that both Vidal Rodeiro and Nádas' work on modularisation and Steinberg and Hyder's work on regulation of standards shed considerable light. The former shows that views regarding 'modular qualifications bad, linear good' – and vice versa – do not reflect the complex ways in which modular and linear qualifications operate in schools and colleges. Policy makers', educators', commentators' and the public's perceptions of the value and application of modular and linear qualifications need to be informed by the realities disclosed by sound research. And hence to regulation – the sum of the analysis presented suggests that the current commitments to regulation which attempts to steer and shape, rather than directly control, all aspects of qualification arrangements will be more possible in a context where awarding bodies undertake comprehensive and incisive research and validation work – leading to a well-balanced set of system arrangements.

**Tim Oates** *Group Director, Assessment Research and Development*

## Editorial

The articles in this issue illustrate the challenges facing the assessment community across a range of areas. In Issue 8 of *Research Matters* Johnson and Nádas reported on their study of on-screen marking and the impact of mode on reliability and marking behaviours. They concluded that further research should be conducted to explore mode-related effects in marking essays of greater length. Shiell *et al.* here report on their latest study where they replicated the research in the context of extended Advanced GCE essays. In their work on the effects of GCSE modularisation Vidal Rodeiro and Nádas combined quantitative and qualitative research methods to investigate the impact of modular assessment on GCSE students, specifically in the key area of *flexible assessment*. This allows units to be taken at the end of a course in a linear fashion or to be taken in different sessions throughout the course to follow a more unitised approach to teaching and learning. Shaw and Crisp then discuss how perceptions of validity have changed over time and the issues that have led to these changes. Their work illustrates the complexity of validity and its importance given the high stakes nature of educational outcomes, their uses and the inferences based upon them.

Black and Gill test the hypothesis that Critical Thinking (CT) skills are transferable and can be applied to other subjects in a beneficial way. They consider some of the research evidence in this field and discuss the best way to deliver CT so as to foster transferrable CT skills and dispositions. Novaković and Greatorex focus on a review of literature, theory and method in their article on the context of vocational qualifications. They consider how instruments used in the vocational field could be used to compare different types of qualifications and the effectiveness of existing methodologies. This is a complex area, fuelled by expectations that standards should remain constant over time, across subjects, between awarding bodies and between task and test demands.

Bramley and Oates describe two research methods that are used within Cambridge Assessment both for operational and experimental purposes. Rank ordering and paired comparison methodologies have been used extensively in the comparability research work of Cambridge Assessment and their use in operational aspects of examinations is being explored and validated. Steinberg and Hyder discuss the need for minimal and useful regulation and how new patterns of engagement between those concerned with the creation and use of assessments can lead to the better regulation of public examinations.

**Sylvia Green** *Director of Research*