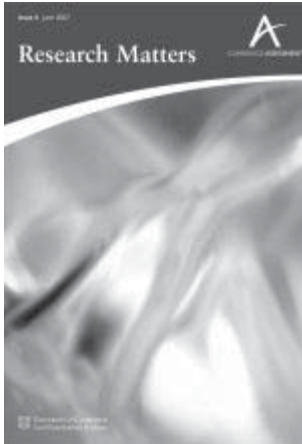


# Research Matters : 4

A CAMBRIDGE ASSESSMENT PUBLICATION



- 1 **Foreword** : Tim Oates
- 1 **Editorial** : Sylvia Green
- 2 **The 'Marking Expertise' projects: Empirical investigations of some popular assumptions** : Dr Irenka Suto and Rita Nadas
- 6 **Did examiners' marking strategies change as they marked more scripts?** : Dr Jackie Greatorex
- 13 **Researching the judgement processes involved in A-level marking** : Victoria Crisp
- 18 **Quality control of examination marking** : John F. Bell, Tom Bramley, Mark J. A. Claessen and Nicholas Raikes
- 22 **Quantifying marker agreement: terminology, statistics and issues** : Tom Bramley
- 28 **Agreement between outcomes from different double marking models** : Carmen L. Vidal Rodeiro
- 34 **Item-level examiner agreement** : Nicholas Raikes and Alf Massey
- 37 **Fostering communities of practice in examining** : Andrew Watts
- 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  
The full issue and previous issues are available on our website:  
[www.cambridgeassessment.org.uk/research](http://www.cambridgeassessment.org.uk/research)

## Foreword

As new technologies begin to emerge in assessment, it sometimes feels as if progress is being made on delivery mechanisms without commensurate development in understanding of measurement and related aspects of assessment processes. This fourth edition of *Research Matters* should be a corrective to that. It drills down deeply into marking. It extends understanding, through empirical work on marking processes; but in addition validates the research methods which can be deployed to undertake such work. This work is not only vital for developing a better understanding of contemporary assessment processes and for developing more refined procedures regarding management and quality assurance of marking, it is vital also for benchmarking the changes which will inevitably come with the introduction of new technologies. There is a misplaced debate on 'will we use technology to do what we do now, only more efficiently, or will it be used to assess in quite new ways?' All the evidence we have to date suggests that even if we set out to use new technology simply to make existing processes more efficient, the introduction of new technologies – such as on-screen marking – always has some impact on the technical characteristics of assessment. The work outlined in this issue challenges assumptions and generates new evidence on marking; in doing this it additionally provides us with an invaluable reference point as we monitor the impact of structural change.

*Tim Oates Group Director, Assessment Research and Development*

## Editorial

The main themes of this issue relate to the psychology of marking, cognitive processes affecting accuracy, and issues related to quality assurance in marking processes. The first three articles focus on marking practices and the factors that impact on them. In their article Suto and Nadas report on their work in the context of marking expertise, considering the demands and expertise that the marking process entails. Greatorex, in her work on examiners' marking strategies, investigates how cognitive strategies change as examiners become more familiar with mark schemes and candidates' answers. In the third article on cognitive strategies, Crisp explores the judgement processes involved in A-level marking for both short answer questions and essays. The next four articles explore quality control of marking processes. Bell *et al.* outline new opportunities for quality control systems given the development of new technologies. Bramley offers a review of the terminology used to describe indicators of marker agreement and discusses some of the different statistics which are used in analyses. He goes on to consider issues involved in choosing an appropriate indicator and its associated statistic. In her study on double marking Rodeiro evaluates the agreement between marks from double marking models and discusses the advantages and disadvantages of blind and non-blind marking. The fourth article on quality assurance from Raikes and Massey focuses on the extent to which different examiners agree at item level and how far this agreement varies according to the nature of the item. This article contributes to the debate on the way in which new item level data, available due to advances in technology, could and should be used in future quality assurance procedures. In the final article Watts discusses the importance of fostering communities of practice amongst examiners in the context of new and developing technological systems and procedures.

*Sylvia Green Director of Research*