

What makes some GCSE examination questions harder to mark than others? An exploration of question features related to marking accuracy

Irenka Suto and Rita Nadas

**Paper presented at the British Educational Research Association Conference,
London, 5-8 September, 2007**

Background

It has long been established that marking accuracy in public examinations varies considerably among subjects and markers (Murphy, 1982; Newton, 1996). This is unsurprising, given the diverse cognitive strategies that the marking process can entail (Suto and Greatorex, *in press*), but what is it about some questions that makes them harder to mark accurately than others? It is to be expected that multiple-choice questions with single letter responses are easier to mark than long essays. However, there exist variations in accuracy within subjects, among questions with much more similar formats and structures. Are there distinct but more subtle features of questions and their mark schemes that can affect accuracy? Such features could potentially contribute to a broad rationale for designating questions to markers according to personal expertise.

Aim

The main aim of this study of GCSE maths and physics examination marking was to identify question features that can distinguish those questions that are marked highly accurately from those that are marked less accurately.

Methods

The study comprised an exploration of maths and physics questions from past GCSE examinations, which were marked in an experimental setting by groups of markers and yielded differing marking accuracies. The questions also varied in their difficulty for GCSE candidates, and in the cognitive strategies needed to mark them.

Kelly's Repertory Grid technique and semi-structured interview schedules were then used in meetings with highly experienced Principal Examiners, who had led the experimental marking of the questions. The data generated comprised ratings for each question on a number of question features (constructs). The ratings were analysed together with the marking accuracy data, enabling an investigation of possible relationships between each question feature and (i) marking accuracy, (ii) question difficulty for the candidate, and (iii) apparent cognitive marking strategy usage.

Findings

For both subjects, marking accuracy was found to be related to various subject-specific question features, some of which were also related to question difficulty (for the candidate) and/or apparent marking strategy complexity. For example, for maths, questions with follow-through marks were marked less accurately than those without follow-through marks; for physics, questions with lots to read were marked less accurately than those with little to read. For both maths and physics, several other subject-specific question features identified by the Principal Examiners were found to be unrelated to accuracy.

Discussion

Overall, the findings have potential implications for the management of markers and for question design. Where differences in accuracy among individual markers, or among marker types are found, a number of subject-specific question features could

be used to decide who should mark what. However, there is a need for further research into the generalisability of the study's findings, involving other examinations and subjects.

References

Murphy, R. J. L. (1982) A further report of investigations into the reliability of marking GCE examinations, *British Journal of Educational Psychology*, 52, 58-63.

Newton, P. (1996) The reliability of marking General Certificate of Secondary Education scripts: mathematics and English, *British Educational Research Journal*, 22, 404-420.

Suto, W.M.I. and Greatorex, J. (*in press*) What goes through an examiner's mind? Using verbal protocols to gain insights into the GCSE marking process, *British Educational Research Journal*.