What goes through a marker's mind? Gaining theoretical insights into the A-level and GCSE marking process

Jackie Greatorex and W. M. Irenka Suto (Presenters and Convenors)
Research Division
Cambridge Assessment

A report of a discussion group at Association for Educational Assessment - Europe, Dublin, November 2005.

Address for correspondence
Dr Jackie Greatorex
Research Division
Cambridge Assessment
1 Hills Road
Cambridge
CB1 2EU
E-mail: suto.i@cambridgeassessment.org.uk
Telephone: 01223 553855
Fax: 01223 552700

www.cambridgeassessment.org.uk

UNIVERSITY *of* CAMBRIDGE
Local Examinations Syndicate

Cambridge Assessment

Cambridge Assessment is the new name for the University of Cambridge Local Examinations Syndicate, a not-for-profit, non-teaching department of the University of Cambridge. As Europe's largest assessment agency, Cambridge Assessment is highly influential in the development and delivery of assessment throughout the world.

What goes through a marker's mind? Gaining theoretical insights into the A-level and GCSE marking process

This is a report of the discussion group convened for the AEA-Europe Annual Conference on 03/11/05 by Dr Jackie Greatorex and Dr Irenka Suto from Cambridge Assessment.

The discussion group was structured into three presentations, each of which was followed by discussion. The presentations were:-

1. Identification of the cognitive strategies used to mark GCSEs (JG and IS),
2. A quantitative analysis of marking strategy usage (IS),
3. A validation of the strategies: their usage in on-screen A-level and GCSE marking (JG).

1. Identification of the cognitive strategies used to mark GCSEs

Background
The question of what goes through a marker's mind has been studied from a psychological perspective previously, in the contexts of university examinations and the testing of English for Speakers of Other Languages. It is surprising, however, that, although the experiences of GCSE and A-level examination markers have been studied from other social science perspectives, a cognitive psychological approach has not yet been applied.

General Certificates of Secondary Education (GCSEs) are obtained by most 16 year olds in England at the end of their compulsory schooling. GCSE courses can be studied in many different subjects, for example, mathematics, English, sciences, design and technology. Typically, a GCSE is assessed primarily through examinations, which are marked by examiners who do not know the candidates. However, there is also a teacher-assessed aspect to the qualification.

Aims
- To investigate the cognitive strategies used when marking GCSEs.
- To interpret them within the context of psychological theories of human judgment.

Methods
Business studies and mathematics GCSEs were chosen for use in this study because they had contrasting mark schemes. 6 business studies and 6 mathematics examiners marked 10 scripts at home then received feedback on their marking. Subsequently, the business studies examiners marked 46 business studies scripts at home, and the mathematics examiners marked 40 mathematics scripts at home. All examiners marked the same scripts. After marking, the examiners visited the Research Division to mark up to 5 scripts whilst 'thinking aloud' and were then interviewed about marking.

The resulting verbal protocols and the interviews were transcribed. All the transcripts were read through, and consequently, a coding frame was established. The transcripts from marking whilst thinking aloud were divided into small sections, each referring to what each examiner said about each question part in each script. For each section of transcript, two researchers judged independently which cognitive strategy the examiner was using. Any discrepancies were discussed between the researchers until agreement was reached. If the coding frame had been found to be

inappropriate for the data, then another coding frame would have been developed at this stage. The coding frame was also applied by one researcher to the post-verbal protocol interview transcripts, and some of these transcripts were checked by the other researcher.

Findings
Five cognitive strategies were identified.

Matching is a cognitive process that has been investigated in psychological research in other contexts (for example, Baddeley, 1999). An examiner can use the matching strategy when the answer to a question is a visually recognisable item or pattern, for example, a letter or part of a diagram. The examiner looks at a particular location in the answer space and judges whether the candidate's answer in that space matches the mark scheme answer.

Scanning is a cognitive process that is well-established in psychological studies of attention (Kramer *et al*., 1996). Examiners use it when they survey the whole of the answer space designated to a question to find whether a particular detail in the mark scheme is in the candidate's answer. This detail could be simple, for example a letter or part of a diagram. Alternatively, it could be more complex, for example, a point in an argument; in such cases, further cognitive marking strategies might be also be used.

When *evaluating*, an examiner pays attention to either all or part of the answer space for a question, and the candidate's answer is processed semantically. The examiner awards marks, bearing in mind the structure, clarity, factual accuracy and logic or other characteristics of the candidate's answer given in the mark scheme.

Scrutinising follows on from, or is used together with, other cognitive strategies. It is used only when a candidate's response is unpredicted. The examiner tries to establish whether the candidate has given a valid alternative to the answer in the mark scheme. To do this, the examiner evaluates numerous features of the candidate's response with the overall aim of reconstructing the candidate's line of reasoning or establishing what the candidate has attempted to do.

When a candidate has written nothing in the answer space, the examiner looks at the space once or more and then gives 0 marks. This is the *no response* strategy.

Full details of this research are given in Suto and Greatorex (in submission, a).

2. A quantitative analysis of marking strategy usage

Aims
- To investigate the contexts and relative frequencies with which the five cognitive strategies are used.
- To identify any relationships between cognitive strategy usage and marking reliability.

Methods

From the coding of the verbal protocol data, the frequencies of cognitive strategy usage on each question were quantified for each individual examiner, as well as for all examiners in each subject.

The reliability of marking of individual examiners was calculated using the marking of the 46 business studies and the 40 mathematics scripts (which were marked silently). For each examiner, these 'experimental' marks were compared with (i) the marks awarded when the same scripts were marked professionally, the previous year; and (ii) the Principal Examiner's 'experimental' marks. Individual question parts on which significant differences in marking occurred were also identified.

Findings

There were some differences in strategy usage among individual examiners. However, the most prominent differences were between subjects and among questions. For most questions, examiners used strategy combinations rather than single strategies.

Within each subject, no clear relationships between strategy usage and marking reliability were found, suggesting multiple successful ways of marking some questions.

Full details of this research are given in Suto and Greatorex (in submission, b).

3. A validation of the strategies: their usage in on-screen A-level and GCSE marking

Background

Generally, A-levels are taken by 18 year olds in England. Universities and employers use A-levels and other qualifications (amongst other information) as indicators of whether a young person is suitable for employment or further study. Young people tend to take A-levels in three or four different subjects; they are mostly assessed through examinations marked by examiners who do not know the candidates.

Aim

To further validate the five cognitive marking strategies identified previously (Suto and Greatorex, in submission, a) by confirming their usage in other marking contexts.

Methods

Verbal protocol data that had originally been collected for other purposes from GCSE mathematics and A-level physics examination markers were reanalysed qualitatively. These were collected when the 10 experienced mathematics and physics examiners (professional teachers or retired teachers) and 4 novice markers (mostly mathematics and physics graduates) were 'thinking aloud' whilst marking on screen.

Findings
The analysis revealed that the strategies are used (i) in A-level as well as in GCSE marking; (ii) by inexperienced as well as by experienced markers; and (iii) when marking scripts on screen rather than in the traditional paper format.

Delegate discussions
Following each of the three presentations, lively discussion ensued around the following themes:

Research method ('thinking aloud' whilst marking)
Some delegates questioned the various assumptions upon which this psychological method of investigation is based. It had been assumed that: (i) the marking process will be similar in both a professional (or 'live') marking context and an experimental context; (ii) examiners will be able to make their marking processes explicit whilst marking; and (iii) examiners will recognise their marking strategies when they are described subsequently by a researcher.

Paradigm loyalties
The delegates were split in their loyalties towards the different paradigms adopted by social scientists in their attempts to understand human behaviour. In the most simplistic sense, this might be seen as a divide between sociological and cognitive psychological perspectives. Some delegates referred to work on professional expertise by Michael Eraut, Christian Heath and David Greatbatch; among these authors ethno-methodological principles have been followed rather than those of cognitive psychology.

Models of mind
Additionally, the delegates were divided in their commitments to different models and concepts of 'mind'; this split was connected to the paradigm loyalties outlined above. For example, it was pointed out that socio-culturalists would claim that 'mind' does not reside 'in the head' and that we create labels (in this case marking strategies) by measuring them.

Responses to the above challenges
With regard to the research method, the presenters stood by the assumptions that they had made. Indeed, when the research findings were discussed with senior examiners, the assumptions were regarded as reasonable and the cognitive strategies were validated further - they received strong recognition from all the examiners. Furthermore, in the related fields researchers hold the view that 'thinking aloud' whilst marking slows down cognitive processes but does not change them significantly.

In response to some delegates' arguments over the adoption of a cognitive psychological paradigm and alternative models of mind, one delegate suggested that individuals with concerns should return to conference next year to present their own thoughts on the question of "What goes through a marker's mind?". Another delegate asked, "If this is not what goes through a marker's mind, then what does?" Such a discussion might advance knowledge in the field and address the latter delegate's question.

It is also worth noting that educational measurement or assessment (the context of the conference) is predominantly about describing and measuring constructs such as mathematical ability/ potential / achievement / performance. To reject the ideological stance of the research presented in the discussion is arguably also to reject the notion of educational assessment or measurement.

Delegates and presenters alike agreed that to undertake research, one must make a reasoned commitment to one paradigm or another.

Uses of the research
Some delegates agreed with the presenters that the strategies gave markers a language with which to communicate with one another about what they do and that the strategies will be useful in the training of new examiners. Others recognised the strategies and thought they might be used in other marking contexts.

Suggestions for further research
As mentioned above, the researchers did not claim that their model of five strategies was exhaustive, describing all aspects of GCSE and A-level marking. The examination papers included in the research tended to involve short answers, i.e. questions worth up to 10 marks. Together with delegates, they discussed the possibility that the marking of longer essay questions might involve the 'evaluating' strategy but in that context it may well need to be divided into sub-strategies. Further research could investigate this speculation. Indeed it is planned that the Research Division of Cambridge Assessment might undertake some research on the cognitive strategies used to mark questions with longer answers. A further suggestion for possible additional analysis was the question of whether marks awarded whilst thinking aloud are comparable to those awarded whilst marking silently.

References

Baddeley, A. (1999) *Essentials of human memory* (Hove, Psychology Press).

Kramer, A., Coles, M. & Logan, G. (1996) *Converging operations in the study of visual selective attention* (Washington DC, American Psychological Association).

Suto, I. and Greatorex, J. (*in submission, a*) What goes through an examiner's mind? Using verbal protocols to gain insights into the GCSE marking process*.

Suto, I. and Greatorex, J. (*in submission, b*) A quantitative analysis of cognitive strategy usage in the marking of two GCSE examinations.