



UNIVERSITY *of* CAMBRIDGE
Local Examinations Syndicate

On Screen Marking of Scanned Paper Scripts

A report published by the University of Cambridge Local Examinations Syndicate

Monday, 7 January 2002

© UCLES, 2002

What is On Screen Marking of Scanned Paper Scripts?

On screen marking of scanned paper scripts is the process by which paper scripts are scanned (at UCLES or at regional bureaus) and the images transmitted to an image server at UCLES. Copies of these images are then distributed electronically and marked on screen by examiners. Question-level marks and examiners' annotations are captured by UCLES throughout the marking process, without manual intervention.

What do we want it for?

The system allows UCLES to enhance service to our customers in a number of ways:

- **Faster and more flexible marking and script management processes.**
Two factors will work together to help us ensure that scripts are marked in the shortest time possible.
 - Improved monitoring lets us track the status of every script throughout the process, allowing tighter management and rapid identification of bottlenecks.
 - Dynamic apportionment of scripts to offsite examiners – scripts are only allocated and distributed to an examiner when he or she is ready for them, ensuring that no examiner is waiting for scripts while there are scripts waiting to be marked.
- **More efficient quality control.**
We can more effectively monitor examiner reliability while marking is underway, allowing us to identify and investigate problems at an early stage, when interventions can be made most efficiently. A conceptual diagram of a prototype workflow for marker standardisation and quality control for a typical examination is attached as Appendix A.
- **More detailed feedback to centres about how their students performed.**
Since every candidate's mark on every question is recorded, we may provide information to centres about how well their candidates did on each question compared with other candidates. This may help centres identify strengths and weaknesses in their coverage of a syllabus. Some individual candidates (for example, those considering re-taking an examination) might also be interested in similar information about their individual performances.
- **Smooth transition to online assessment**
Our customers are likely to want a mixture of paper-based and computerised examinations for some time. The paper scanning approach allows us to use the same basic infrastructure to process both types of assessment, facilitating a smooth transition and allowing us maximum flexibility to meet our customers' needs.

Does it work?

UCLES' first major test of on screen marking of scanned paper scripts was conducted in winter 2000 (though several small scale trials were conducted before this). The main aims of this study were to prove the practical possibility of the scanning and electronic movement of scripts, to provide research data concerning examiner reliability, and to uncover issues, both technical and human, to be investigated in later stages of the development of a production system.

Three November 2000 components were chosen, one each from O Level Mathematics, A Level Geography and A Level English Literature. These components were selected because they covered a variety of question types and script types. Scripts were scanned in Cambridge. After scanning, the paper originals were distributed and marked conventionally¹, and the images were distributed over the Internet to examiners' homes for on screen marking. Two types of on screen

¹ All scripts were marked conventionally, and these were the marks used operationally, since the reliability of on screen marking had not been proven.

marking were investigated: whole script marking, where examiners marked all of a script, and individual question marking, where scripts were split by question and examiners specialised in particular questions². Five examiners for each subject took part in the study, all chosen for their proficiency as examiners and **not** for their IT literacy. They all marked the relevant component operationally (on paper) immediately before doing their screen marking³.

An annotated screenshot of the marking software used by examiners is attached as Appendix B.

In general, the scanning process implemented for the trial worked and the practical viability of scanning scripts and distributing them electronically was demonstrated. Many opportunities for improvement were identified in an issues register, and this has informed subsequent developments. Examiners' impressions and suggestions were captured in a questionnaire, the main findings of which were:

- Examiners could generally schedule downloads such that they had work available when they were ready to mark;
- Most scripts were at least as legible on screen as on paper;
- Most examiners felt that they ended up with the same marks on screen as on paper. Of the three who didn't, two marked English Literature and one marked Geography. The main concerns related to the trial nature of the facilities provided in the prototype software for annotations and scrolling, and to a perceived need to know a script's total mark;
- Mathematics examiners tended to find question marking boring or less rewarding than marking whole scripts, and some English Literature examiners said that they needed to mark a whole script to award a fair mark;
- All the examiners who took part in the study would mark on screen again.

Examiners' responses to the questionnaire also contained many valuable insights and suggestions for improvements.

Analysis of marks data suggested that:

- For the Mathematics component, examiners applied similar standards and were similarly consistent across the three marking methods (on paper, whole scripts on screen, and individual question marking on screen);
- For Geography, although most of the marking was satisfactory, one examiner was a little more severe when marking on screen and one examiner, whose paper based marking was reasonably consistent, was inconsistent when marking on screen;
- For English Literature, two examiners were a little more severe on screen than on paper (it made little difference whether they were marking whole scripts on screen or individual questions). Examiners tended to be most consistent when marking on paper and least consistent when marking question apportionments (on screen). This may have been because examiners were unable to be influenced by a candidate's performance on other questions when the scripts were split by question.

The results indicated that with suitable modifications to the software used by examiners, screen based marking of whole scanned paper scripts would be likely to be as reliable as conventional marking. Individual question marking required more investigation, particularly for English Literature.

² All three components were marked on screen as whole scripts, but only the Mathematics and English Literature components were marked as individual questions.

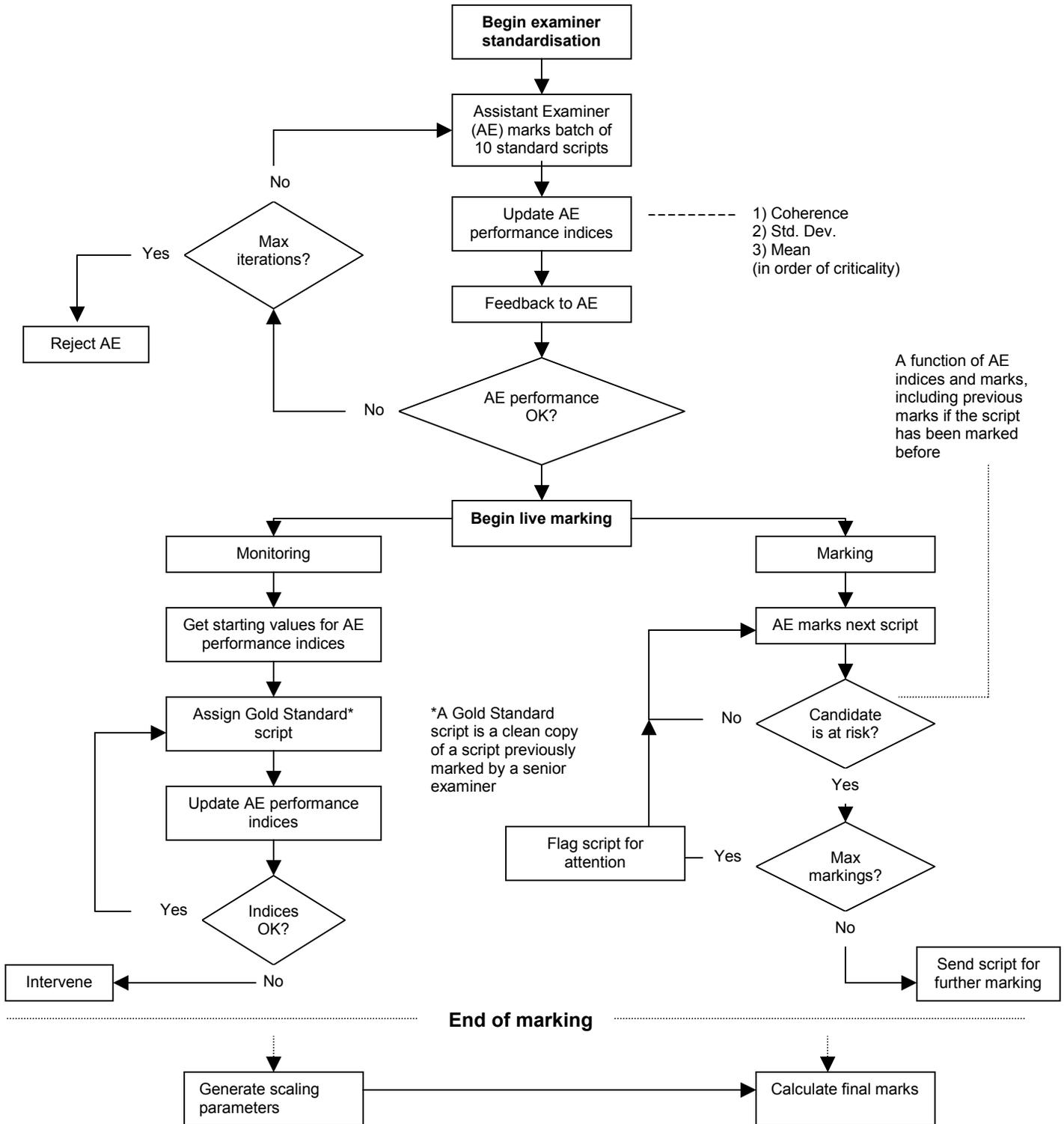
³ As far as possible, individual examiners only marked a particular script once (i.e. either on paper, or in one of the screen marking modes).

What next?

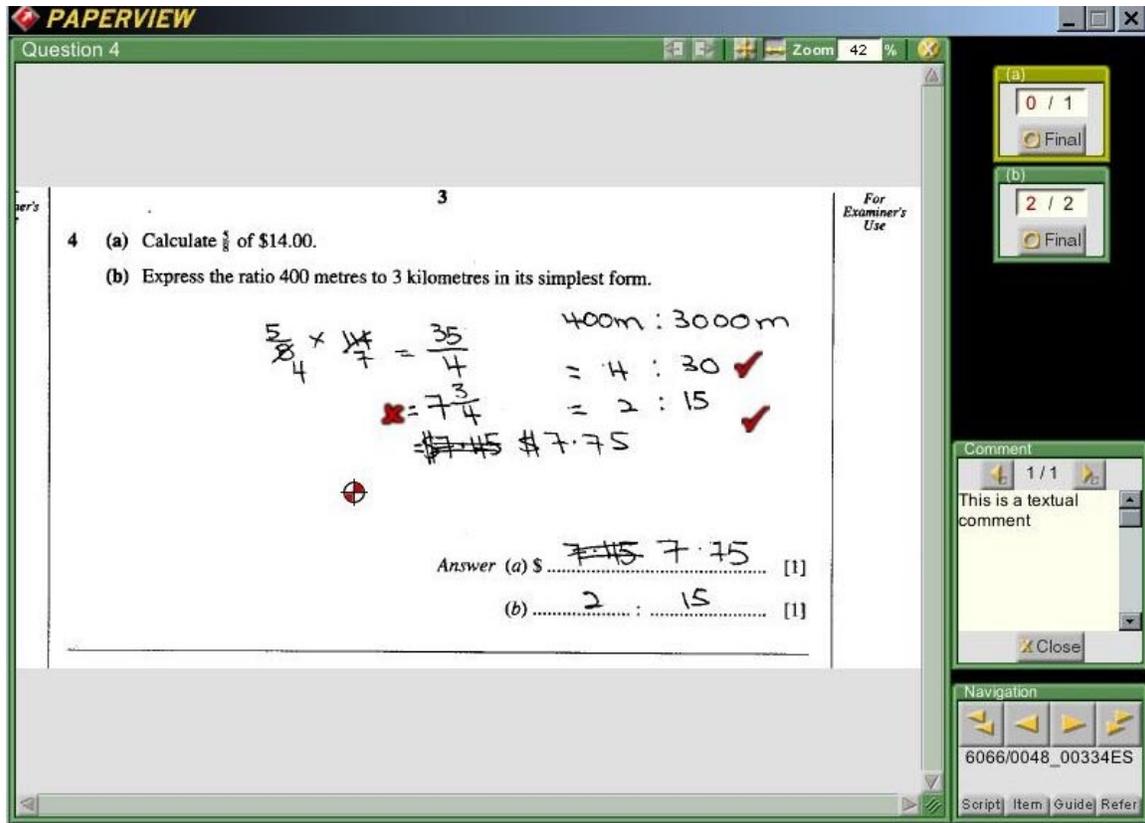
UCLES is engaged in a programme of research and development to identify the refinements needed for a production quality system and the contexts in which screen based marking is fully valid and reliable. Contracts for the development of production software have been entered into, and we expect to introduce and ramp up the use of on screen marking of scanned paper scripts over the next few years.

This summary was prepared by Nicholas Raikes, Interactive Technologies for Assessment and Learning Unit, University of Cambridge Local Examinations Syndicate. Many people contributed to the work reported.

Appendix A:
Conceptual diagram of a prototype workflow for examiner standardisation and quality control for a typical examination



Appendix B:
Screenshot of the prototype marking software used by
examiners



Main window (left hand part of the screen)

This displays the candidate's answer and any ticks and crosses entered by the examiner. The partially shaded circle is the anchor for the examiner-entered textual comment displayed on the right hand side of the screen. The controls in the top right of the window are used for setting display properties and for navigating between pages of multi-page answers.

Navigation box (bottom right of screen)

The outer buttons with the double triangles are used for navigating to the same question in the next or previous script. The inner buttons with the single triangles are used for navigating to the next or previous question in the current script.

The Script button at the left of the bottom row of buttons displays a list of the scripts in the current batch, together with an indication of which ones have been completely marked.

The Item button contains a list of the items in the current script, again with an indication of which ones have been completely marked.

The Guide button displays guidance notes for the examiner on how to mark the currently displayed question.

Examiners use the Refer button to refer a problem script (with a comment) to an administrator or Team Leader.