



ASSESSMENT and
QUALIFICATIONS
ALLIANCE

**A feasibility study on anonymised marking
in large-scale public examinations**

Jo-Anne Baird & Neil Bridle

CONTENTS

SUMMARY	II
CENTRES' VIEWS.....	II
EXAMINERS' VIEWS.....	II
ADMINISTRATION	II
GENERAL	III
INTRODUCTION	1
PREVIOUS RESEARCH	1
SYSTEMS OF ANONYMISED MARKING	3
DESIGN OF THE STUDY	5
EXAMINATION PAPERS	5
ADMINISTRATIVE MODIFICATIONS	5
EVALUATION.....	6
RESULTS & DISCUSSION	6
CENTRE QUESTIONNAIRES	6
EXAMINER QUESTIONNAIRE	9
ADMINISTRATION	11
CONCLUSIONS.....	14
REFERENCES.....	14
APPENDIX A – ANONYMISED EXAMINATION BOOKLET	16
APPENDIX B – ADDITIONAL ANSWER SHEET	17
APPENDIX C – INSTRUCTIONS TO INVIGILATORS	18
APPENDIX D – DISCREPANCY FORM	20
APPENDIX E – EXAMINER INSTRUCTIONS.....	21

SUMMARY

Names of centres were removed from summer 2000 GCSE and A level examiners' stationery as part of a Department for Education and Employment initiative. Awarding bodies expressed concern about the possibility of mis-matching candidates to their examination work if candidates' names were not visible on the examination script. A pilot study on anonymised marking was conducted in the summer 2000 AQA (SEG) GCSE English (2400PF) foundation tier examination. The purpose of this study was primarily to investigate whether administrative errors would occur due to the concealing of candidates' names. For two written response question papers, candidates used examination booklets on which they wrote their name in the top right hand corner and then concealed it from the examiner by folding and sticking the corner down. Over 34 000 candidates (68 000 scripts) were included in the study.

Centres' views

Over seventy percent of centres responded to a questionnaire designed to elicit centres' views on the pilot study. Many centres wrote that they had to spend extra time checking candidates' centre and candidate numbers because their names were not visible on the scripts. Sticking the flap down on the booklet was problematical because the glue was not sticky enough and because candidates were reluctant to lick the scripts in a number of centres. Licking the scripts to stick them down was seen as unhygienic. A self-sealing booklet would circumvent this problem, but may be costly. Under the guidance of invigilators, candidates generally found the instructions for completion of the examination scripts straightforward. Surprisingly, only 42% of examination officers who responded to the questionnaire expressed support for the introduction of anonymised marking, 31% were against it and 21% were unsure¹.

Examiners' views

Only 119 (36%) examiners responded to the questionnaire eliciting examiners' views of the pilot study. Problems in identifying candidates were mentioned by several examiners, but there were few cases per examiner. Many examiners welcomed the system, but there was a higher proportion of negative than positive comments on the questionnaire. The perceived insinuation of biased marking was seen as offensive by many examiners: a view that was also prevalent in responses to the centres' questionnaire. Respondents questioned the evidence for biased marking. Additionally, examiners were concerned that marking was de-personalised in an anonymised system.

Administration

Few administrative problems arose due to the introduction of the system. The AQA Mark Processing Department was able to process candidates that had not been registered for the examination (pirate entries) without much additional difficulty. Very few flaps had to be

¹ Five percent of examination officers did not answer this question.

opened by examiners, to check candidates' identities. There were also very few errors (less than one percent) in the centre and candidate numbers on the examination booklets. Flaps on a sample of scripts (approximately 10%) were opened and no mis-matching of candidates with scripts was found.

Although concealing centre names was a blanket requirement and not strictly part of the study, data were collected on how successfully centre names were concealed. In one third of centre packs, the centre name could be identified by means such as postage franking. Checks on the scripts themselves revealed that candidates' names were not properly concealed on six percent of scripts. Concealing of centre and candidate names may be dealt with more effectively in future years, as centres become more accustomed to the procedures.

General

Concealing candidates' names from examiners is not a panacea for marking bias, as handwriting style and the content and style of the language used reveal personal characteristics of the candidates. Further, at times it may be necessary for the valid examination of a particular subject to set questions that require candidates to reveal personal characteristics. In one of the questions in the current study, candidates were required to write a letter on behalf of their school. Many candidates used the real name of their school and signed the letter with their own names. This was a valid question to assess letter-writing skills at foundation level. Since the aim of the pilot study was to investigate administrative problems, the inclusion of this question did not interfere with the study's objectives, but it does serve to highlight that under an anonymised marking system examiners would have to consider whether it was necessary for candidates to reveal personal characteristics to answer the questions.

There was a general feeling amongst centres, examiners and examining board staff that the anonymised system introduced more *possibilities* for administrative errors. Findings of this research indicate that these administrative errors were surmountable.

The findings of this pilot study suggest that with some modifications the introduction of anonymised marking could be successful in an administrative sense, albeit with cost implications. However, support for the system may not be very forthcoming from teachers (and examiners), who frequently saw the system as another attack on teachers' professional integrity. Teachers may not have considered that anonymised marking could act as protection against allegations of marking bias directed at individual examiners or the awarding bodies in general.

INTRODUCTION

Personal characteristics of examination candidates, such as their ethnic origin, sex and socio-economic status should not be relevant to the assessment of their work for public examinations. Candidates and centres are identifiable by name to the examiners who mark the candidates' work in public examinations in England and Wales. This procedure means that there is a *possibility* that bias due to personal characteristics of the candidates could be exhibited in marking. Awarding bodies try to ensure that no bias is present in the marking by prescribing what candidates must do in the examinations to be awarded marks. This information is publicly available, following the examinations, in published marking schemes. Of course, it is recognised that professional judgment of examiners is also a necessary ingredient for accurate marking, so marking schemes alone may not eradicate bias. A highly pragmatic reason for examiners not exhibiting bias in marking is that they work under a great deal of time pressure. Many examiners say that they simply do not have time to focus upon the names of the candidates or the school when they are marking. Nonetheless, there is a possibility that bias (conscious or otherwise) could be present in the marking.

The Department for Education and Employment asked the awarding bodies in England and Wales to remove centre names from all examination stationery in the summer 2000 GCSE and GCE examinations. Matching candidates' work to the correct candidate is clearly crucial for accurate reporting of results. The additional removal of candidates' names from scripts could make the matching process error-prone. Therefore, AQA decided to conduct a pilot study on the concealment of candidates' names on the examination scripts. This paper reports the results of that anonymised marking pilot study which was conducted on GCSE English foundation tier papers in summer 2000. A brief review of research on anonymised marking is considered below, before the study itself is presented.

Previous research

Anonymised marking has been introduced in many UK universities, despite a lack of research on its effects. A study conducted by Bradley (1984) in higher education provided some evidence that anonymised marking could eliminate sex bias in undergraduate project results. The projects were marked by a project supervisor who knew the student and then double-marked. Second examiners were less familiar with the students, but did know the students' names and therefore their genders. Bradley's rationale was that second examiners would show more sex bias because they had less information about the students and there was therefore more scope for stereotypes to operate. Only projects in which there was disagreement about the marks were included in the study. In four university departments that used this double-marking procedure, second examiners marked men's projects more extremely than they did women's. However, in a polytechnic department where second examiners were unaware of students' names, and generally therefore their genders, no such pattern in the marks was found. Two conclusions were drawn from these results. First, the author claims that greater

knowledge of the student will, by itself, reduce sex bias. In this regard, Bradley's paper may be too optimistic. Some summary data giving overall marks for projects for males and females from first and second examiners would have been useful, since males may have gained better marks from the first examiners too. The second conclusion from this study is that anonymised marking eliminated gender bias.

Newstead and Dennis (1990) replicated Bradley's (1984) study and did not find the pattern of results that Bradley interpreted as biased marking. Where there were disagreements between examiners, men were more likely to have their marks raised than were women, but this difference did not reach statistical significance. In a further study (Newstead and Dennis, 1990), a comparison was made of marks in institutions using anonymised and non-anonymised marking, but no effect of marking procedure was demonstrated.

A study by the Scottish Examining Board (1992) investigated the possibility of marker bias in Standard Grade English and History, although the effects of anonymised marking were not investigated in this study. As part of a controlled experimental design, examiners were sent scripts that varied in terms of the achievement record of the presenting centre (high or low), the handwriting on the script (neat, messy or typed), gender of the candidate and ethnic origin of the candidate (Asian or British). The gender of the examiner was also taken into consideration. The only significant effect found in the English scripts was that typewritten scripts scored a lower mark than hand-written scripts. This was thought to be due to the lack of use of the spell checking facility in word processing the typewritten script. For the History scripts, those attributed to females were awarded more marks than those attributed to males. The results of this study are suggestive that, in at least one subject, females may gain from gender bias in marking. Perhaps girls were evaluated more highly because History is often assessed by essays and girls are thought to be better at extended writing (Punter and Burchell, 1996).

The only published study that investigated the effects of anonymised marking in public examinations was conducted by Baird (1998). The research investigated whether gender bias was present in the marking and if so, whether anonymised marking could eradicate it. Although examiners may not look at the candidates' names, it was thought that it may be possible to guess the sex of the candidate from the style of the handwriting. A pilot study found that judges could guess the sex of a writer correctly on 75% of presentations. Examiners were sent different forms of the same work, which varied in terms of whether the script was anonymised, had 'male' or 'female' writing and male or female names (on non-anonymised scripts). English Literature and Chemistry examiners participated in the study. No bias was found in the marking and, therefore, there was no effect of the anonymised marking procedure. Of course, it is not possible to demonstrate with such studies that no bias in marking exists, but this study did indicate that there was no bias in favour of either sex.

Recognition memory is highly accurate and examiners can often recognise a script that they marked previously, even when it was originally marked amongst an allocation of hundreds of scripts. Consequently, asking the same examiner to mark differently presented, but identical, scripts is not methodologically sound. Studies of marking bias can therefore only address average differences in marks due to experimental manipulations, such as changing the name on the script to a different sex. By its nature, bias affects individuals. Regrettably, research cannot tackle these individual instances of bias because there is not enough experimental control: it is always possible that bias was not present at all and that some other aspect of the work in the script was not favoured by the marker. Of course, if the marking is not in accordance with the marking scheme, then it is clear that there is a discrepancy. In an individual case this discrepancy cannot be attributed to bias, unless there is a consistent pattern across cases.

Systems of anonymised marking

At the very least, anonymised marking involves concealing candidates' names from the examiner. It can also involve trying to conceal the identity of the centre through which the candidate entered the examination. Candidates and centres are typically represented by numbers instead of names in anonymised marking systems. Of course, this is not a fool-proof system, as examiners come to know things about the candidate that the system is designed to prevent them knowing (such as the numbers associated with particular students in university departments: Watson, 1997). Handwriting may also give clues to gender or ethnic origin. As discussed above, Baird (1998) found that people could identify the gender of candidates from their handwriting style alone with an accuracy rate of 75%. Style of writing could give even more away about the candidate. Further, in some examination papers, the content of candidates' writing may be highly associated with their personal characteristics. For all of these reasons, anonymised marking is not necessarily sufficient to eliminate bias in marking, should bias be present in examiners.

However, even if anonymised marking is not sufficient to prevent the operation of bias in marking, it could be argued that it is one of the necessary conditions for reducing bias. In other words, if it could overcome any bias at all, then it is worth implementing. However, in an examination system with millions of candidates' scripts, the name of the candidate can be very useful in ensuring that errors are not made in the processing of the results. In effect, to introduce anonymised marking could have the consequence of random error in candidates' results, if scripts are mis-allocated to candidates. The system may also be more open to abuse, since it could be more difficult for centres to check that the correct candidate has sat the examination. In an effort to prevent impostor candidates, current inter-board examination procedures require candidates to sign booklets, as well as to print their names.

There are a number of ways in which anonymised marking can be implemented. Possibilities include;

- a flap with the candidate's name on it that is folded back or stuck down,
- pre-printed number labels (to reduce the frequency of transcription of candidate and centre number errors),
- bar-coded labels and
- a label over the candidate's name that can be peeled off, if necessary.

Each of the above has advantages and disadvantages in terms of manageability and cost of the systems, for examination boards and centres alike. Anonymised marking could introduce considerable additional administrative burdens on centres. For this reason, the views of centres were sought in the current study.

Candidates currently write a centre number and a candidate number on the front of their examination scripts. These numbers could be used to match the work to candidates if no name was present on the script. Indeed, the numbers are one of the two main ways in which scripts are currently matched. The Northern Ireland Council for the Curriculum, Examinations and Assessment (NICCEA) has had a system of anonymised marking in place for decades. Their system conceals both the centre and candidate names from examiners and numbers only are used to identify candidates. Entry figures for examinations in Northern Ireland are much smaller than those in some English examinations, which makes a system of anonymised marking more manageable. NICCEA ensure that scripts are correctly matched to candidates by collating the scripts at the board and rectifying any problems before they are distributed to examiners. The size of the entries for syllabuses run by some of the English boards means that adopting the NICCEA system could significantly delay the processing of results as well as increase the costs.

Another possible problem with a system in which numbers only are used to identify candidates is that mis-matching of examination work to candidates could occur. An internal Associated Examining Board Report (Baird, 1998) investigated the frequency of errors made by candidates in writing their centre and candidate numbers. At A level and GCSE less than 1% of scripts had centre or candidate number errors. Centre number errors should be easy to correct, as scripts are sent to examiners in centre packs. Only 0.5% of scripts had candidate number errors. However, in a system of anonymised marking this would produce a large number of scripts to be matched to candidates. Further, awarding bodies receive a number of scripts each year in which the candidate has not been registered for the examination. Thus, the findings of this, earlier, study indicated that a system that retained candidates' names would be preferable. Additional reasons of cost and administrative efficiency led us to trial a system in which the candidates print and sign their names on the corner of the examination scripts. The corner of each script was then stuck down before the scripts reached the examiner. An example front page of an examination booklet can be found in Appendix A – Anonymised examination booklet.

The main objective of the current research was to investigate the administrative problems (if any) that could arise, for the awarding body and for centres, due to the introduction of anonymised marking using this method.

DESIGN OF THE STUDY

Examination papers

Two written papers in the AQA (SEG) GCSE English foundation tier (2400PF) examination were selected for the study. Centres were notified that the study would take place in this examination. Over 34 000 candidates were entered for the foundation tier papers. As such, the papers provided a test of administrative issues that could arise due to large scale examining. It was thought that errors would be more likely to occur in GCSE foundation tier papers than in higher tier GCSE or A level examinations. The study was designed to target potential problems, since the objective was to find out what administrative problems anonymised marking could incur, so that AQA could prepare to deal with these issues, should anonymised marking be introduced.

Administrative modifications

Centre names were removed from stationery for every A level and GCSE examination in summer 2000, due to the DfEE's request to all awarding bodies. So for the purposes of this study, only candidate names needed to be removed from or concealed on examination stationery. Examination answer booklets were modified, as reported above (Appendix A – Anonymised examination booklet). Additional answer sheets were also provided (Appendix B – Additional answer sheet).

Attendance records were also adapted. AQA attendance records usually show candidates' names on all three copies. For the study, the attendance records showed candidates' names on the top copy only, which the centre was asked to retain and to send the second, anonymised, copy to the examiner.² In standard examining procedures, invigilators send the top copy of the register to the examiner, so a special set of invigilation instructions was sent to centres (Appendix C – Instructions to invigilators). The instructions covered issues such as checking candidate identity and ensuring that candidates knew how to deal with the new style of examination booklet.

Examiners were notified at the co-ordination meeting that an anonymised procedure would be used. They were asked to record cases in which the examination booklets were not sealed, or had to be opened, on a new form (Appendix D – Discrepancy form). Instructions about how

² Due to an IT department error, standard attendance records were sent to centres prior to the anonymised attendance records. The standard attendance records were erroneously used in 41% of centres for paper 5 and 54% of centres for paper 6.

to deal with problem cases (such as two candidates with the same numbers) were also given (Appendix E – Examiner instructions).

Evaluation

- Questionnaires were sent to centres and to examiners, eliciting their views on anonymised marking in general and whether they had any administrative problems. Half of the centres received questionnaires following paper 5 (2400/5) and the others received questionnaires following paper 6 (2400/6).
- Frequency of centre and candidate number errors was investigated.
- Numbers of and reasons for open booklets were explored.
- A sample of booklets was opened, to check for mis-matching of candidates to their work.
- Qualitative data were collected on the problems encountered within Marks Processing Department, due to the anonymised marking procedure: especially in relation to candidates who had not been registered for the examination.
- Whether centre names could be identified from the packs of scripts was investigated.

RESULTS & DISCUSSION

Centre questionnaires

Seven hundred and thirty two centres entered candidates for the foundation tier of the examination and of those, five hundred and twenty four responded to the questionnaire (72%). It should have been no more difficult for centres to check the attendance register because the top copy was identical to the standard register. Approximately three-quarters of centres that responded said that it was just as easy to check candidates' attendance with the anonymised system (Table 1).

Comments noted that if a candidate did not appear on the register (for example if he/she was changing tier), the procedure is that the centre writes the candidate's details onto the register by hand. The examiner's carbon copy below then showed the candidates' names. It may be difficult to circumvent this problem in an operational system, as matching the candidate with their work is a higher priority than anonymity. Another problem was that the register paper slipped when invigilators wrote on it, which meant that the carbon copy was not very accurate due to the close proximity of lines on the register and this slippage. These problems with the design of the attendance register are not peculiar to this study, but they are more important in this context because the examiner received the second and not the top copy of the register.

Centres' most common response was that it was just as easy to verify candidates' identities and to package the scripts into envelopes as in previous years. However, a sizeable minority felt that these administrative tasks were more difficult than in previous years. The impact of the anonymised system upon centres will depend upon their internal systems for checking

attendance and candidates' identities, so it is not surprising that centres responded differently to these questions. From centres' comments, it was obvious that many examination officers spend a lot of time checking candidate and centre numbers. This explains why so few scripts have centre and candidate number errors when they arrive at the examination board (Baird, 1998 and the section on administration below). Many centres (151 out of 732) wrote that candidates had written their numbers wrongly and rectifying this was made more difficult and time-consuming by the anonymised booklet or that they were concerned that candidates could write their numbers wrongly. One centre commented that,

“The new format phased candidates who are already in a high state of emotion. The Foundation Level chosen were perhaps more affected by this. Many missed the candidate number or put it incorrectly - these had to be checked from the seating plan. All in all it doubled ... exam checking time”.

Table 1 The new anonymised scripts made it easier/just as easy/more difficult to carry out each of the following tasks.³ (% of centres)

	easier	just as easy	more difficult
i) check candidates' attendance	1%	74%	24%
ii) verify candidates' identities	1%	71%	26%
iii) package scripts into envelopes	2%	64%	31%

Centres also encountered problems with candidates either not sticking the flaps down at all, or sticking them down prematurely, prior to the attendance check (7). Some centres wrote that they dealt with identification problems by uncovering the name and re-sealing the flap. Comments suggest that there was a general problem with the glue used on the booklets and that it was difficult to stick the flap down (53). A technical solution to this will have to be found if the system is to be introduced, since the flap should seal effectively and it should be obvious if it has been tampered with once it has been sealed.

Further problems associated with the type of anonymised booklet used in the study were encountered by centres. Candidates often did not want to lick the glue (50 responses), complaining that it was unhygienic. A religious issue also arose, as one candidate was concerned that the glue may have contained pig derivatives. Another candidate was allergic to the adhesive. One examinations officer also complained about hygiene problems, saying that he/she did not want to have to open flaps that had been licked, to identify candidates. Worse still, the examinations officer did not want to lick the glue again to stick it back down. A self-sealing flap would be more hygienic and may be more technically adequate, but its cost would have to be explored.

³ Percentages do not sum to 100 due to missing responses in each of the tables.

Size of centre appeared to have an effect upon attitudes to the anonymised system, as small centres often commented that they had no problems due to the number of candidates they had entered for the examination. Northern Irish centres had very few problems with the system, commenting that they were used to dealing with anonymised marking (albeit a different system) in the NICCEA board. A number of centres commented that any problems they experienced were largely because only one examination was being dealt with in this way. This meant that the invigilators had to use different procedures, which was complicated by the sharing of the examination hall by candidates entering for more than one examination. The new procedures had to be explained to the candidates. One centre commented as follows:

“The time taken to explain it all to the cand. was annoying, particularly when you have to tell them all the other details of the exam. Resulted in exam being 10 mins. late starting - detracted from the start of the exam by testing cand. patience - exam start should be straightforward, easy and confidence building”.

Another centre suggested that it would have been useful to have had the answer booklets earlier, so that the administration could have been explained to candidates prior to the examination. It was suggested that the answer booklets could have been packaged separately from the question papers, to facilitate familiarisation with the new design before the examination. Example copies of the new style of booklet were sent to centres prior to the study, but it may be that these examples did not reach the invigilators. Filling in the booklets was not seen as more problematical than filling in the standard booklets, so problems in explaining the administration to candidates is only a teething problem, which would be overcome with experience of the new booklet design.

Centres were often unsure about how to deal with candidates who had special consideration in the examination. Centres allowed these candidates to be identified by name in the normal way. Amended procedures for special consideration of candidates will have to be introduced if anonymised marking becomes operational.

One question on paper 6 required candidates to write a letter. Frequently candidates identified themselves and their school in these letters. Questions like this are valid examination questions in an English examination, despite anonymity being compromised. Since the purpose of the study was to uncover administrative difficulties with anonymised marking, the inclusion of this question does not compromise the study. However, it does highlight the fact that if an anonymised marking system were to be introduced more generally, examiners would need to consider carefully whether it would be necessary to compromise candidates' anonymity to answer the question paper. To test letter writing skills at this level, the answer is likely to be yes.

Two thirds of centres who responded ticked that candidates found the instructions on the scripts straightforward (Table 2). Invigilators' role in this was emphasised in the comments on this question. A minority of students stuck the flap down backwards, rather than on the front of the script.

Table 2 Did candidates find the instructions on the scripts straightforward? (% of centres)

Yes	No	Don't know
67%	18%	12%

When examinations officers were asked whether they were in favour of anonymised marking, the most popular response was 'Yes' (Table 3). However, support for anonymised marking was still a minority view. Teachers often dismissed the idea that there was bias in the marking, questioning the evidence for biased marking. For example, one examination officer wrote "Is there any evidence to suggest that it makes any difference to how a marker assesses a paper? Is there any research to date on this idea, or is it just in its infancy?" The small number of studies conducted in this area provides little evidence for the existence of bias. Comments noted that examiners do not have enough time to look at the names on the scripts. Interestingly, centres almost never cite marker bias as the reason for enquiries upon results or appeals against examination results. Overall, marking appeared to be viewed as not biased by respondents, but worryingly, one examinations officer wrote that "As an experienced marker I would argue that anonymised marking could take away from a degree of subjectivity." Professional judgment is required in the assessment of examination work, but should not be based upon information about the centre or candidate. This issue is discussed further below, in relation to the examiner questionnaire.

Table 3 Are you in favour of the introduction of anonymised marking? (% of centres)

Yes	No	Don't know
42%	31%	21%

There were also concerns about mixing up candidates' work either within the centre, or at the board. Earlier notification of the study was requested by several respondents and the need for uniformity of examination procedures across awarding bodies was emphasised. Whether the cost of the introduction of an anonymised marking system will be passed on to schools was questioned.

Examiner questionnaire

One hundred and nineteen examiners responded to the questionnaire (36% of the papers 5 and 6 examiners). The majority of examiners who responded felt that it was either easier or just as

easy to check that the scripts had arrived, mark the scripts and complete the optical mark reader (OMR) sheets with the anonymised system (Table 4). However, approximately one third of examiners wrote that it was more difficult to check that the scripts had arrived, complete the OMR sheets and complete discrepancy forms. Examiners commented that they had no means of checking the candidate to whom the script belonged if the candidate number was wrong. They also pointed out that the name was used as a double-check when they completed the OMR sheet in previous years and there were concerns that there could be more errors due to the absence of the names. (E.g. “Where cand. had been marked absent on the outside of centre sacks, it was not always easy to identify the numbers and match with mark sheets - where you also have the name you have 2 items to check with each other. Teacher's handwriting not always clear.”) The design of the attendance register was also commented upon by examiners: “It would be helpful to alter the layout of the attendance register. For absent cand. there is not enough space which makes checking difficult”.

Table 4 The new, anonymised scripts made it easier/just as easy/more difficult to carry out each of the following tasks. (% of examiners)

	easier	just as easy	more difficult
i) check that all of the scripts had arrived	3%	58%	34%
ii) mark the scripts	1%	87%	8%
iii) complete the OMR mark sheets	5%	55%	36%
iv) complete discrepancy forms ⁴ for individual candidates	4%	45%	34%
v) package scripts into envelopes	3%	86%	8%

The highest number of flaps that had to be opened by an individual examiner was 14. Examiners reported that scripts often had no candidate numbers, candidates were often not listed on the mark sheets and candidates' writing of the numbers was sometimes illegible.

Seventy percent of the examiners who responded wrote a comment at the end of the questionnaire: 44 of the 119 respondents wrote negative comments and only 12 wrote positive comments (the remainder were mixed). Those with strong feelings are more likely to respond to questionnaires and with such a low response rate, the questionnaire results are not likely to be a good indication of examiners' views on anonymised marking generally. Nonetheless, it is clear from the results of the questionnaire that at least 13% of all the examiners involved had negative comments about the pilot study. Many examiners welcomed the system and saw it as not causing problems.

⁴ These forms register any problems, such as if the script did not have any centre or candidate numbers.

“Although it made some of the admin/checking more difficult - it didn't add any great time to the task. Also it meant you had to double check numbers rather than rely on a name - which in itself is a good thing. Primarily I think it is an excellent idea - it guarantees fairness regarding gender/race. However 'perfect' we think we are as markers and that we mark each paper fairly - the anonymity guarantees this and one doesn't fall into making any assumptions. Definitely the way ahead for all exams, at all levels - inc. KS2/3.”

Conversely, some (5) examiners wrote that they were offended by the implication that their marking could be biased. Lack of political faith in teachers has de-moralised the profession in recent years and this was seen as another indication of the lack of trust. The two comments below illustrate these points.

“I am 100% against this system. As ... [a senior examiner] I have an enormous amount of reviewing of ... [scripts] to do and the sealed corners with the totals in the middle of the page mean that you can't flick through a centre, but must turn each ... [script] individually. This is IRRITATING & TIME CONSUMING. In the foundation tier in particular, students frequently make a complete mess of their cand. numbers and you can end up with the same cand. number on three ... [scripts]. With the attendance registers blanked out too, it is then IMPOSSIBLE to find out whose the scripts are. In the past, you could check the names easily. I also object to the move on principle. It is INTERESTING to see names and adds a touch of humanity to a fairly dreary and soul destroying job of prime marking. It is insulting to professional people to take the names away as it suggests they can't be trusted not to be biased.”

“This procedure causes a waste of time ... In shortness of allocated time I find it hard to believe that any marker has time to brood over the lists of cand. names to the extent that prejudice could occur. In fact I find it insulting that AQA or any other body should suppose that markers are prejudiced. This procedure made me wonder whether I shall mark next year if this is how markers are viewed.”

The de-personalised nature of the process was mentioned by several examiners (12), who felt that it was better to know the name of the candidate to establish a rapport with them. For example, one examiner wrote “To have personal insight into a cand. is a good thing - it is too anonymous now”. However, it is unclear whether this rapport translates to more accurate marking.

Administration

Six percent of flaps were found to be open in each paper when the scripts were returned by the examiner to the board. In each paper, approximately fifty scripts were reported as being opened by the examiners on the Marking Discrepancy Sheet. The most common reasons for

examiners having to open the flaps were that the candidate was not listed on the marksheet or the wrong candidate number had been used. Examiners reported a few instances in which candidates had used the same numbers. It is possible that examiners could have under-reported the number of scripts they had to open, since they had to complete a new form to do this. However, the figures available suggest as many as three in four of the flaps that were open had been left open by the candidates. Responses to the centre questionnaire suggest the most likely causes for candidates failing to stick the flaps down were lack of stickiness of the glue and refusal to lick the script. These findings suggest that examiners would not have to open many flaps in an operational system: an average of less than one per examiner was reported as opened by the examiners.

Mis-matching of candidates with their examination work was one of the main concerns in this study. AQA's Marks Processing Department logged the number of errors candidates made in writing their centre and candidate numbers. Findings concerning the frequency of centre and candidate number errors were in line with previous research (Baird, 1998). Less than 1% of scripts had errors in the centre and candidate numbers (0.6%: see Table 5) and only 0.2% had errors in the candidate numbers. In contrast to the previous study (Baird, 1998), centre number errors were mainly detected at the board, rather than by the examiners. Anonymisation of centres probably made it more difficult for examiners to find and correct centre number errors. As previously thought, incorrect centre or candidate number errors are infrequent.

Table 5 Frequency of errors in the candidate and centre numbers on the scripts

Paper	Centre no. errors detected by		Candidate no. errors detected by		Total
	examiner	board	examiner	board	
2400/5	22	131	54	33	240
2400/6	17	85	31	28	161
Total	39	216	85	61	401

Note: 34 235 candidates entered each paper in the examination.

For a random sample of centres, flaps were opened following the examination period, as another check that scripts had been correctly matched to candidates. Approximately ten percent of centres (and candidates) were sampled in each paper. Details on very few (25 out of 7454) of the opened scripts differed from the registered numbers and names. A few scripts had no numbers (2) or names (4) and in some cases candidates had spelt their names slightly differently (14). In three cases, the candidates had omitted either their first name or their middle name on the script. Either the surname or the first name on the script differed from the details held on the database in the remaining three cases. No case of a mis-matched script was found in the sample.

Although anonymising centre names was not part of this study, data were collected regarding the extent to which centres could be identified when the packs of scripts were returned to the board. In each question paper, centre names could be identified in approximately one third of packs by franking on the envelope, inclusion of a compliments slip, completion of the centre name box on the front of the envelope, a special delivery postage sticker, a letter from the centre or special arrangements forms. In one case, the invigilator wrote the centre name on the attendance register. This figure is likely to be an under-representation of the number who could be identified by examiners, as examiners will have removed letters from centres that needed to be forwarded to the Subject Officer.⁵ It appears that the anonymising of centre names was not very successful in its first year of introduction. As centres become accustomed to the practice, it may operate more effectively. However, centres did write in the questionnaire that not using their franking machines for examination scripts would be problematical for them. Forms for the special consideration of candidates should be sent directly to the board by centres, but many included them in the pack of scripts to examiners. In view of centres' practice, forms for special consideration of candidates should be revised to anonymise the centre names.

One specific concern raised prior to the study was that scripts received for candidates that had not been entered for the examination would be difficult to deal with in an anonymised marking system. However, no particularly problematic cases arose in the pilot study. The AQA Marks Processing Department was able to handle the pilot study effectively, but there were concerns that errors could arise due to the system. This department has a very heavy workload throughout the examination and enquiries upon results seasons and the anonymised marking pilot was an additional burden. Many centres rang the department to ask how they should deal with the anonymised scripts. Indeed, the department also fielded a large number of calls regarding anonymising centre names although this was a blanket requirement and not directly related to the study. One centre that took part in the study did not put candidate names on any of their examination stationery for other subjects.

Although the anonymised scripts did not cause much additional work in operational procedures during the summer, considerable extra work has been caused by the anonymised scripts in relation to enquiries upon results. For cases in which results are pending, the flaps on each of the scripts in the centre have to be opened to check whether a candidate's scripts are present. As the pilot was conducted in only one subject, the additional workload was manageable this year, but could be excessive if the pilot was extended. Flaps also have to be opened on the scripts as part of the clerical checks procedures for enquiries upon results (this has been put into action for GCSE English this year). The enquiries upon results season is

⁵ Of course centres could often be identified by students' letters written in answer to one of the questions on paper 6, but this method of identifying the centre will not be common to other subjects and has therefore been excluded from this analysis.

already a very pressured time and the need to open script flaps in all subjects would exacerbate these pressures considerably. Given the extension this year of arrangements for giving candidates access to their scripts, it seems likely that additional processing staff would be required if anonymised marking was introduced.

CONCLUSIONS

Few administrative difficulties were encountered within the board as a result of anonymising candidates' scripts. Some unexpected administrative difficulties were identified in the centre questionnaire and a few problems were encountered by examiners. Support for the system was not very forthcoming from teachers and examiners, but it is not known whether students and parents would be supportive of anonymised marking. It would be useful to extend the evaluation of anonymised marking to investigate candidates' and parents' views of the procedure.

The findings of this pilot study suggest that with some modifications the introduction of anonymised marking could be successful in an administrative sense, albeit with cost implications. The following issues need to be addressed if fully anonymised marking is to be introduced.

1. A technical solution is needed for the slippage on the attendance registers.
2. A hygienic and cost-effective examination booklet has to be found.
3. Forms for special consideration of candidates need revision, so that at least centre names are concealed.
4. Centres need to be advised of the change as early as possible (e.g. before Easter).
5. Instructions to centres should be emphasised by being sent to the subject departments and the examinations officers at each centre.
6. Examiners and centres must be advised that a system of anonymised marking serves to protect examiners from possible allegations of biased marking.

REFERENCES

- Baird, J. (1998) What's in a name? Experiments with anonymised marking in A level examinations. *Educational Research*, 40, 2, 191-202.
- Baird, J. (1998) *Feasibility of anonymised marking in GCSE and A level examinations*. AEB internal report RAC/791.
- Bradley, C. (1984). Sex bias in the evaluation of students. *British Journal of Social Psychology*, 23, 147-153.
- Newstead, S.E. & Dennis, I. (1990). Anonymised marking and sex bias in student assessment. *Assessment and Evaluation in Higher Education*, 15 (2), 132-139, Summer.

- Punter, A. & Burchell, H. (1996). Gender issues in GCSE English Assessment. *British Journal of Curriculum & Assessment*, 6 (2), 20-23.
- SEB (1992). *Investigation into the effects of the characteristics of candidates and presenting centres on possible marker bias*. Scottish Examination Board internal report.
- Watson, R. (1997) On your marks for changes in the exam system. *Times Higher Education Supplement*, London.

October 2000



**ANONYMISED MARKING:
Instructions for invigilators for components
2400/5 & 2400/6 – Summer 2000**

The government has encouraged awarding bodies to consider introducing a system of anonymised marking in GCSE and A level examinations. On behalf of the awarding bodies, the Assessment and Qualification Alliance (AQA) is piloting an anonymised marking system on the above SEG GCSE English Foundation Tier components. A copy of the front page of the new style of anonymised examination answer booklet is attached. The information in the top, right hand corner of the examination answer booklet is concealed from examiners by folding and sticking the top, right flap of the answer booklet. **Please use only the new style of examination answer booklets and supplementary answer sheets provided for these components.**

The following instructions for invigilators for components 2400/5 and 2400/6 in Summer 2000 supplement the *GCE and GCSE Awarding Bodies Instructions for the Conduct of Examinations 2000*.

Excerpt from standard instructions	Additional notes for anonymised examination answer booklets
<p>7. Identification of candidates 7.1 The invigilators must be satisfied as to the identity of every candidate attending each examination session. The Head of Centre must ensure that appropriate arrangements are in place to enable all invigilators to carry out adequate checks on the identity of candidates.</p>	<p>The identity of each candidate must be checked before candidates seal the examination answer booklets.</p>
<p>10. Starting the examination 10.2 (i) Instruct candidates to enter their names, centre number, candidate numbers and the component code/paper details on their answer booklets where applicable and complete other details as required;</p>	<p>Bring the new style of examination answer booklet to candidates' attention. Instruct them to complete the front of the booklet and not to seal the flap until the end of the examination, when instructed to do so by the invigilator.</p>

./..

Excerpt from standard instructions	Additional notes for anonymised examination answer booklets
<p>13. Completing the attendance register 13.1 The invigilator must complete the Attendance Register during the examination, in accordance with the instructions of the Awarding Body.</p>	<p>Note the new style of Attendance Register: after marking the register, you should retain the top copy (showing candidate names) and forward the bottom copy (without candidate names) to the examiner.</p>
<p>18. Collection of scripts Arrangements must be made for all scripts to be collected, placed in the order shown on the Attendance Register and checked to ensure that all scripts are present and that candidates have used their correct centre and candidate numbers. <i>(Centres are advised to ensure that all scripts are collected before candidates are allowed to leave the examination room.)</i></p>	<p>Checking that candidates have written the correct centre and candidate numbers and put their candidate name on the fold-over flap is crucial in the anonymised marking system, as examiners will identify candidates by these numbers only.</p>

If you have any queries about invigilating the 2400/5 and 2400/6 examinations, please telephone Neil Bridle at AQA (01483 477634)

Appendix D- Discrepancy form



GCSE English Anonymised Marking Discrepancy Sheet

SUMMER 2000

Component Code:

Please list below, all the candidates for whom you have removed the gummed down flap on the script. Please refer to the instruction leaflet for details of when this action is applicable.

Centre Number	Candidate Number	Candidate Name	Reason

Please return this form to Mr N Bridle, AQA, Stag Hill House, Guildford, Surrey, GU2 7XJ in the pre-paid envelope supplied.

Signed:

Person No:

Assessment and Qualifications Alliance is an alliance of AEB/SEG, C&G and NEAB and is a company limited by guarantee registered in England 3644723 and registered charity number 1073334. Registered address Addleshaw Booth & Co., Sovereign House, PO Box 8, Sovereign Street, Leeds LS1 1HQ. Kathleen Tattersall Director General.

ANONYMISED MARKING

Instruction Leaflet for Examiners Marking Components 2400/5 & 2400/6

- A new anonymised marking procedure is being piloted by AQA in the above components, on behalf of all of the awarding bodies. There is no research evidence for the presence of bias in marking, but concealing candidates' names protects examiners from accusations of bias associated with candidates' or centres' names. You are therefore asked to mark your scripts this summer without knowing the names of the candidates or the centres concerned.
- To ensure anonymity, candidates' personal details are to be hidden on the examination script, under the gummed flap on the right hand corner. Do not open this flap unless absolutely necessary. The conditions under which you should open the flap are given below.
- The instruction in the Examiner's Handbook (which will be included in the examiner stationery pack) should be adhered to, although the following instructions supersede those in the handbook.
- Please mark all scripts for a centre prior to transferring the marks to the Examiner Assessment Mark Sheets (EMS). Check that the centre numbers are recorded on the scripts and, if not please enter them.
- Check that the candidate numbers are recorded on the scripts and, if not please open the gummed down flap on the top right hand corner of the script and record the candidate details on the anonymised marking discrepancy form (M/DISC: enclosed with this despatch).
- Once you have marked all the scripts within a centre you must transfer the marks to the EMS (see instructions for completing Examiner Mark Sheets – part of stationery pack).
- Where you receive a script for a candidate not shown on the Examiner Mark Sheet, the centre, candidate number and mark should be recorded on Form EX/SUPP. **You will not be able to record the candidate name on the EMS.**
- If you have two scripts with the same candidate number on them you should open the gummed down flap on the top right hand corner of the scripts and record the candidate details on the anonymised marking discrepancy form (M/DISC). **Please leave the EMS blank and do not complete form EX/SUPP for these candidates.**
- If you have any queries concerning whether you should open the flap on the answer booklets please telephone Neil Bridle at AQA 01483 477634. We will ask for your experiences of this new procedure in a questionnaire following the marking period.