A QUALITATIVE EXPLORATION OF KEY STAKEHOLDERS' PERCEPTIONS AND OPINIONS OF AWARDING BODY MARKING PROCEDURES

Rachel Taylor

Abstract

Past research into perceptions of GCSEs and A-levels suggests that how informed individuals are of current marking procedures, and the level of information stakeholders receive varies considerably. However, ensuring certain individuals associated with the examination process sufficiently understand awarding body procedures, including those adopted for marking examination papers is highly important. A lack of understanding by stakeholders suggests awarding bodies are not transparent enough, which can result in misconceptions and mistrust of their work. Additionally, ensuring individuals understand the examination system fully might lead to greater confidence in the system, something which is important to awarding bodies given the competitive environment in which they operate. Previous research into perceptions of the examination system has failed to explore individual's perceptions of the actual processes adopted for marking examination papers, or opinions on what these procedures should be. Therefore, the main aim of this study was to address this gap in the literature, by exploring perceptions and opinions of marking procedures among key stakeholder groups. Fourteen semi-structured interviews were conducted with teachers, parents, and examiners and a focus group was conducted with five GCE students. The findings suggest that parents, teachers and students are largely unaware of current marking procedures. Additionally, there also appears to be a lack of understanding of the examination system in general. Considering these findings, it is suggested that AQA should do more to increase the transparency of their routine processes and should aim to increase understanding of the examination system amongst key stakeholder groups.

Introduction

Each year students sit over 3.5 million GCSE and GCE examinations with AQA. Ensuring these students receive the right result on their examinations is of high importance, not only to the students themselves, but also to centres, teachers and parents. Examination grades impact upon the life chances of candidates, as well as the reputation of teachers and the position their centre occupies in performance league tables. The task facing awarding bodies is to ensure each candidate is given the mark on an examination script that most accurately and fairly reflects that particular script's merit, and to adopt a fair and reliable system to assigning a candidate's mark. In particular, the mark awarded to a candidate's work should be independent of the examiner doing the marking.

UK awarding bodies currently adopt a hierarchical approach to assigning a candidate's mark on an examination paper, based on a hierarchical system of examiners. Usually, only one examiner marks each script, and a candidate's mark is determined by the standard set by the

AQA

most senior examiner. This contrasts with the classical test theory (Spearman 1904a, 1904b) which defines a candidate's 'true' mark as the mean mark awarded by an infinite number of examiners. This can be viewed as a consensual approach to determining the true mark. Therefore, the more examiners that mark a candidate's script, the closer the mark would become to a candidate's 'true' mark. The hierarchical and consensus approaches to determining a candidate's 'true' mark are likely to result in different true marks, impact upon the reliability of marking and have different implications for quality control of marking and re-marking procedures (Meadows & Baird, 2006).

Although widely accepted practice, the hierarchical approach to determining a candidate's 'true' mark may not be a theoretical ideal but more of a practical necessity. In the UK there are approximately 20 million GCSE and GCE scripts marked annually, in addition to the moderation of centre assessed coursework. This highlights the enormous task facing awarding bodies every year. Considering the shortages of examiners under the current system it is difficult to entertain the possibility of adopting an approach where more than one examiner marked every examination script. The financial implications of using even double marking are also evident, in terms of the extra costs of examiners and the additional resources required. Furthermore, recent research shows only small gains in marking reliability through the use of double marking (Fearnley, 2005). These gains are likely to be outweighed by the obvious practical difficulties of adopting such an approach.

While current marking procedures are well established and arguably practically necessary, how informed key stakeholders such as teachers, students and parents are of these procedures may vary considerably. Little previous research has addressed this issue, despite the seeming importance of ensuring certain stakeholders in the examination system are sufficiently informed about operational procedures. A lack of available information on awarding body operations may suggest that awarding bodies are not transparent enough. Following the production of the Tomlinson report (2002) exploring the circumstances surrounding the A-level results crisis in 2002, AQA stated,

> *"The principles of accountability and transparency … underpin all AQA's procedures, including those which lead to the award of grades. AQA will take all necessary steps to ensure that its procedures are better explained and communicated to the students and centres whose interests AQA serves"* (AQA spokesperson, 2002)

This implies that if AQA's customers do not feel sufficiently informed of marking procedures, AQA are not putting into practice their principles of transparency and accountability as stated. The notion of increased transparency by awarding bodies is supported by Newton (2005), who suggests that, *"Assessment agencies should proactively respond to the challenge of increased transparency"*. According to Newton, increased transparency will lead to long term gains in public trust.

Past research has also suggested that the more informed individuals are about a particular service or organisation the more favourable they tend to be towards it (QCA 2006a, 2006b). Although this implies that increased understanding of marking procedures amongst stakeholders may result in higher confidence in the system, the robustness of the evidence upon which this assertion is made is perhaps somewhat questionable. This suggests further evidence needs to be gathered concerning the link between greater understanding and enhanced confidence in organisations.

Although previous studies into stakeholder perceptions of GCSEs and A-levels have been conducted, these have tended to focus on determining whether individuals believe they are sufficiently informed of marking procedures, rather than exploring perceptions of the procedures adopted for marking examination papers or opinions on how examination papers should be marked. The following section summarises the main findings from two such studies and highlights the limitations of these studies. It also provides an overview of the present study, aiming to extend previous findings and fill the gap that currently exists in the literature.

Overview of previous research

Edexcel (2000) conducted an opinion based survey designed to explore perceptions of GCSEs and A-levels amongst parents. A total of 1,503 parents of students who took their GCSEs and A-levels between 1997 and 1999 were surveyed via the telephone. One of the key findings was that 64 *per cent* of those surveyed stated they were interested in receiving more information about how examinations are set and marked. This suggests a lack of information available in the public domain, at least at the time this research was conducted. Arguably however, when individuals are offered something, such as additional information, it is more likely they will agree they would like this than turn it down, regardless of whether they will actually use it or not.

Another interesting finding was that 76 *per cent* of parents surveyed were aware that examination papers can be re-marked. Although this represents a large percentage of the population, it also shows one in four parents surveyed were not aware re-marking is possible, implying a lack of knowledge of the examination system. However, this research by Edexcel did not explore perceptions of the actual procedures adopted for re-marking scripts, despite these being highly significant to both candidates and parents. If a candidate believes they have received an incorrect mark the re-marking procedure should sufficiently resolve this anomaly. This should either result in a change of mark when appropriate, or the mark should remain the same if the original marking was perceived to be fair and reliable. The result of a re-mark is significant to candidates and parents as a change of grade may impact upon a candidate's place at University, and their future life chances.

More recently, research into perceptions of GCSEs and A-levels has been conducted for QCA by a team at MORI (2006a; 2006b). Following the A-level results crisis in 2002 (where pass rates rose by 5 *per cent* from the previous year) the subsequent independent investigation recommended a need for clearer communication on how examination papers are marked and graded. Additionally, it was recommended that it should be ensured that students, teachers and parents have the right levels of information about the examination system. It seemed likely that confidence in the examination system was adversely affected by the events of summer 2002. Consequently, QCA has monitored public opinion of GCSEs and A-levels since through a number of research studies, interspersed with awareness raising exercises aimed at increasing understanding and ultimately confidence in the examination system. Amongst other initiatives, QCA have developed an official guide to the A-level system (QCA, 2005a), and GCSE system (QCA, 2005b), specifically aimed at students.

The most recent wave of research conducted for QCA concerned perceptions of the GCSE and A-level examination system among teachers, parents, students and the general public (QCA, 2006b). A telephone survey was conducted with a representative sample of 506 teachers from schools and colleges across England, during November 2006. Among the key findings were that only a third of A-level (33%) and GCSE (32%) teachers recall seeing their respective guides to the examination system. Furthermore, only four out of ten A-level teachers (41%) recall receiving information on how exams are marked and graded. Although this shows

information concerning the marking and grading of examination scripts is being distributed to centres to a certain extent, there are still large numbers of teachers who do not recall receiving such information. This implies the recommendation made following the events of 2002 for clearer communication on how exam papers are marked and graded is not being fully met. Another key finding from this research showed that less experienced teachers were less likely to recall receiving any sources of information than their more experienced colleagues, suggesting that any information that is received by centres is not distributed sufficiently between staff of varying levels. Problems therefore appear to exist with the dissemination and accessibility of information to certain individuals.

As part of this research, telephone surveys were also conducted with 138 A-level/GCSE students and 324 parents of A-level/GCSE students. Results showed A-level students were more likely to report having enough information concerning how A-levels are marked and graded (61%) than parents of A-level students (34%). Similarly, more GCSE students (57%) believe they have enough information about how GCSEs are marked and graded than parents of GCSE students (40%). This suggests students may be receiving information through their centre, but this information is less accessible to parents. As such, the recommendations to ensure parents and students have the right levels of information about the examination system are not being fully met. This again highlights problems with the dissemination and accessibility of information, particularly for parents.

While the Edexcel and QCA research provides some insight into perceptions of awarding body procedures, there are several limitations. Most importantly, they lack any in-depth exploration of the issues related to the marking of examination papers. This is primarily due to the use of a survey method in both cases, which used mainly closed questions. Furthermore, both studies covered a large subject area relating to many issues associated with GCSEs and A-levels. No research studies were identified solely exploring the issues associated with the marking of examination papers, despite this perhaps being one of the most important parts of the examination system to candidates, parents and teachers, given the significant impact the grades candidates receive have upon the lives of all these individuals.

Considering the previous research that has been conducted and its limitations, this research aims to provide a much more in-depth exploration of the issues associated with the marking of examination papers. This study aims to explore perceptions and opinions relating to four main topics associated with the marking of examination papers; i) perceptions of how a candidate's mark is assigned – and whether this is considered appropriate; ii) perceptions of the reliability of marking; iii) how we might best ensure quality control of marking; and iv) issues surrounding re-marking.

Method

In order to explore perceptions and opinions of marking procedures a total of fourteen semi-structured interviews were conducted with parents, examiners and teachers, and one focus group with five current A-level students during February and March 2007. Interview and focus group schedules were produced separately for each group of participants to take account of the varying degrees of knowledge of marking procedures anticipated within each group. A summary of the main topics for discussion for each group of participants is provided in Appendix A.

Before conducting any of the interviews or the focus group participants were asked to give their permission for the interview/focus group to be recorded and transcribed. Participants were assured all data collected were confidential and would be anonymised in any resultant report, meaning specific individuals and school/colleges would not be identified. Participants were asked to sign a written consent form for this to which all participants agreed, allowing all the interviews and focus group to be recorded by a digital voice recorder.

Five parents of students who are currently studying for their GCE examinations were recruited to participate via a snowball sampling method, where the initial respondent was asked to recommend another. Although a focus group was originally planned with parents, difficulties in recruiting a group of parents meant this was not possible. Therefore, interviews were conducted with five parents, who were all located within the Leicestershire area. Four of these participants were female and one was male. Interviews were conducted in the parents' own homes and lasted approximately 30 minutes each. Semi-structured interviews were used and discussions were based around a similar scenario to that presented to the students, concerning a student who had recently completed their GCE examinations.

Five examiners were recruited via the telephone from a list of AQA General Studies A examiners who live within the Greater Manchester area. Although General Studies was chosen as the subject through which to recruit participants, there are no issues anticipated that will influence the generalisability of the findings to other groups. The same is true for both teachers and candidates. Four of the examiners who participated were assistant examiners and were male. The remaining examiner was a team leader and was female. Interviews were conducted at AQA where possible or at a public venue closer to the examiners' home. In these situations, the researcher was accompanied by another member of the Research department.

Teachers were initially recruited via the telephone from a list of centres offering AQA GCE General Studies within the Greater Manchester area. Unfortunately, several teachers recruited via this method later withdrew for various reasons. Therefore, a convenience sampling method was used to recruit the remaining teachers. Ultimately, four teachers agreed to participate, all of whom currently teach GCSEs, A-levels, or both. Three of these teachers were female and one was male. The interviews took place at the teacher's centre, or if more convenient in their own homes and lasted approximately 30 minutes each. None of the teachers had any previous experience of examining or acted as examinations officers in their centre. The interviews with teachers were also semi-structured and based round a scenario. The teachers were asked additional questions, based on the assumption that they would be more familiar with the examination system than students or parents. These included discussions of the two theoretical approaches to calculating 'true' marks.

The focus group with students was conducted at a Sixth Form college in Greater Manchester. The Head of General Studies at this centre was approached via the telephone to take part, from a list of centres based in the Greater Manchester area with relatively large entries for GCE General Studies. Having agreed to participate, the Head of General Studies at the centre facilitated recruitment of students for the focus group, and selected five students who he believed would contribute fully to discussions. Of the five students who took part three were female and two were male. All students were in the Upper Sixth Form, aged 17-18. The focus group discussions centred upon a scenario presented to participants concerning a student who had recently completed their A-level examinations. The focus group lasted approximately 45 minutes and students left their usual General Studies lesson to take part, meaning no teacher was present. At the end of the focus group all participants were given a £10 voucher for their

participation. This incentive however did not influence motivation to participate or participants' responses; only the Head of General Studies who facilitated recruitment was aware of this incentive before the focus group.

Following completion of each interview and the focus group participants were given the opportunity to make any further comments or ask the researcher any questions. To analyse the findings the focus group and interview recordings were listened to repeatedly. Key points and themes were identified and then transcribed, meaning only part of the conversation between the researcher and participants was transcribed. The transcription was carried out by the researcher who had conducted the interviews and facilitated the focus group. Transcription of the focus group was assisted by notes taken by another member of the Research department, which aided identification of different individual's responses. The findings are reported by the four main topics discussed.

Results

How are exams marked: the hierarchical vs. consensus approach to assigning the 'true' mark
The first part of this research discussed stakeholders' perceptions of how a candidate's mark is, and should be determined. It also discussed the two theoretical approaches to understanding true scores with teachers and examiners.

Overall, parents and students were largely unaware of current marking procedures and their responses were based on assumptions and perceptions of the system rather than specific knowledge. The majority tended to believe more than one examiner marks each examination script. For example, one parent responded, *"I would have thought 2 or 3, certainly 2"* and another, *"I always thought there was one person checking and one person checking on that"*. Similar responses were received from candidates, such as, *"I'd say two and then if there's a big issue a third one"*. Despite this tendency to believe that more than one examiner marks each script, ideas on how a candidate's mark would actually be finalised under these procedures were lacking. Only when prompted did candidates agree that perhaps some sort of discussion took place between examiners, suggesting they believe some sort of consensus approach might be used to assign a candidate's mark. Parents however were largely unsure how a candidate's mark is finalised. For example, one parent responded, *"I have no idea actually how they do that"*. This suggests they way in which a candidate's mark on an examination paper is assigned, and even how many examiners mark each script is not a topic many parents or students have previously considered.

Compared to parents and candidates, teachers appeared to be more aware of how a candidate's mark is assigned, tending to believe only one examiner marks each script. For example, one teacher commented, *"I thought it just went to one examiner"*. Despite this, it was evident uncertainty still existed and like the students and parents, their responses appeared to be based on perceptions rather than actual knowledge of marking procedures. The theoretical approaches to determining a candidate's mark were also discussed with teachers. Following explanation of the hierarchical approach, teachers tended to agree this was a satisfactory approach for assigning a candidate's 'true' mark. One teacher responded, *"Yes I think that's still acceptable"* and another *"I suppose it has to be, in that we have to value experience"*. Despite this, teachers tended to recognise this might not be theoretically ideal, but more a practical imperative.

Discussions of the theoretical approaches to determining a candidate's 'true' mark revealed that examiners also tended to advocate the hierarchical approach to assigning a candidate's mark. They suggested any benefits of a consensus approach would be minimal in comparison to the extra resources required, and commented that a candidate's mark *"wouldn't be any truer"* from the use of a consensus approach. One examiner stated, *"I don't think necessarily at the moment there is any particular advantage to be gained from moving to another system"*. Additionally, they emphasised the extra resources that would be required, including the financial costs of adopting a consensus approach to assigning a candidate's mark. One examiner commented, *"I suspect that the expense would be phenomenal"*.

Discussion was also undertaken with parents, students and teachers regarding how they would like a candidate's mark on an examination paper to be assigned. The students collectively favoured an approach where at least two and possibly three examiners marked each script. One student stated, *"It's just one person's opinion, so you could have two, maybe three actually"*. Parents also tended to favour a system involving multiple markers, in which a discussion between examiners was used to arrive at the final mark. Teachers tended to agree; for example one teacher commented "*I think multiple marking and a consensus approach would give a much more balanced, perhaps fair result in the end"* and another, *"…with unlimited resources would be better if it went again and again"*.

Despite this preference for a consensus approach some, but not all teachers and parents recognised the practical constraints on the examination system, something which the students were not aware of. For example, when discussing the idea of multiple marking one teacher responded *"I do think it's unrealistic with the amount of time that people have got"*. Parents also commented, *"…like all businesses I'm sure there are financial constraints"* and *"If you think about the logistics of it, how many papers there are to mark each year, I think common sense would say you can't have more than one person checking a paper, other than the person checking the person who's done it"*. Therefore, whilst some parents are aware of the practical constraints on the system others do not take account of these. Even those who do recognise these practical difficulties do not appear to fully understand these, believing for example the system could cope with a system of double marking. This implies a lack of understanding of the scale and complexity of the examination system, which is perhaps to be expected.

To summarise, the parents and students in this research appear largely unaware of how a candidate's mark on an examination paper is assigned, tending to assume multiple marking is used. Although the teachers show higher levels of awareness of marking procedures, they are by no means certain of these and many of their responses were based on perceptions and assumptions. This suggests that information regarding marking procedures is not widely available, or is not distributed sufficiently to stakeholders in the examination system. Further findings show individuals would prefer an approach to assigning a candidate's mark that involves multiple markers, possibly with a discussion between examiners to arrive at the final mark. This suggests to a large extent that individuals appear to lack an understanding of the scale of the examination system, and the practical constraints on this system.

The reliability of marking
The second topic for discussion was the reliability of marking, including perceptions and expectations of how reliable individuals think marking is. Discussion was based around a scenario presented to all participants in which two examiners both marked the same examination paper. Participants were asked whether they thought both examiners would give the same mark, and whether there was any room at all for disagreement between examiners.

The examiners tended to believe two examiners marking the same script would give a similar mark, but not necessarily the same mark. One examiner responded, *"I'd expect the level to be there, but within that level then there would probably be a difference of a mark or two"* and another, *"I'd expect them to be close but not necessarily the same"*. Likewise the teachers tended to agree they would expect a similar mark from both examiners, commenting, *"within a few marks, yes, certainly the same band"*. The teachers however acknowledged this may not always be the case. One teacher responded, *"I would hope they would but I know from experience that that's not always the case…so it isn't an exact science"*.

Candidates' responses differed, from those who believed the two examiners would give the same mark, to those who thought it would just be similar. There was an agreement between the candidates however that examiners should be expected to give the 'right' mark every time they mark an examination script, indicating high expectations of marking reliability. The opinion of parents was also divided; while some believed the two examiners would give the same mark, others doubted this would be the case.

These findings suggest that perceived levels of marking reliability differ between individuals and among groups. The teachers and examiners tended to expect the marks of two examiners marking the same script to be similar and within the same band, but acknowledged they could differ by a few marks. Parents and students tended to have higher expectations of marking reliability, where some expected two different examiners to give exactly the same mark to a script. However, despite these high expectations of marking reliability, many individuals believe there is room for disagreement between examiners. This is an encouraging finding as it suggests individuals are to some extent aware that there is room for legitimate professional disagreement in what is, in some cases, a value judgement.

How is quality of marking assured?
The third topic of discussion centred upon how quality of marking should be assured, with reference to paper based marking. During an examination series examiners may mark up to several hundred scripts each, meaning checking the quality of marking on every script is not practically possible. Instead, for paper based marking, ensuring quality of marking is maintained is done via a script sampling process. Following attendance at a standardisation meeting senior examiners re-mark a sample of ten scripts supplied by assistant examiners (in view of the marks and comments of the original examiner). A second sample is re-marked by a senior examiner part way through the marking period, and serves as a further check on the quality of marking. This section of the research discussed perceptions of how quality of marking is maintained among parents, teachers and students, and explored opinions of how individuals believe quality of marking could best be assured. The suitability of the current script sampling system and suggestions on how awarding bodies could better assure quality of marking were discussed with examiners.

Among parents, candidates and teachers there was little knowledge of current procedures for ensuring quality of marking. Some individuals assumed examiners were left to mark their allocation without monitoring, and that no quality control procedures exist. The students' perceptions differed considerably. For example, one student believed examiners marked without any monitoring, assuming the practical constraints of sending scripts between examiners would prevent script checking. The other students disagreed and one commented, *"I'd say it was like every ten scripts"* and another, *"You have to check, because there's obviously gonna be mistakes"*. Amongst teachers there was also uncertainty. One teacher believed examiners were left to 'get on' with their marking and commented, *"I assumed they just*

got on with it", whilst the others tended to believe some sort of checking procedure was in place but were unaware of the specific details of this. One of these teachers responded, *"I don't know. I think it's checked at some point…I don't know if it's continually checked…"*. Perceptions amongst parents also varied considerably. While parents tended to hope some sort of quality control procedures were in place, knowledge of what these actual procedures were was lacking. For example, one parent responded, *"I don't know what the system is nowadays"* and another *"no idea, I don't know how that would work"*. This suggests parents, teachers and students are largely unaware of how quality of marking is assured, indeed whether quality control procedures exist at all.

Discussions on whether the quality of marking of examiners *should* be monitored throughout the marking period revealed that parents, teachers and students all supported the use of a system of quality checking. Parents and teachers tended to agree a system of regular checking should be implemented. For example, one parent responded, *"Regular checks, I would like regular checks"* and a teacher commented, *"I think consistently throughout…a much more consistent regular checking process"*. A system of regular checking would contrast with the current system in which two samples of marking are re-marked by a more senior examiner at set points during the marking process. It would however be in accordance with recommendations from an independent inquiry following the A-level results crisis in 2002, to explore the use of 'on-the-spot' checks of marking consistency (Baker, McGaw & Lord Sutherland of Houndwood, 2002). One of the teachers supporting a system of regular checking suggested this would help ensure examiners marked consistently throughout the whole marking period, thus alleviating fears the marking of examiners may fluctuate at times. They commented,

> *"…with the best will in the world…you've got 600 scripts to mark and by the time you've got to the 60th of the day you're perhaps not gonna be so accurate and attentive as you were with the first…and the fact examiners know that they've had their checking process and they've just got to get through them now…".*

Generally, the examiners supported the current script sampling system, believing it is sufficient to ensure quality of marking, at least with reference to their own subject of General Studies. They commented that any alternatives to this approach would be difficult to implement, given the practical constraints on the system. In line with ideas offered by parents and teachers, one examiner suggested a regular check on scripts could be adopted. They commented, *"You could take smaller samples as people were marking, just to make sure there was a consistency there"*. They also recognised the benefits a more random approach to checking scripts could offer, by requiring examiners to send randomly selected scripts to the senior examiner rather than those chosen by the assistant examiner. The examiner stated,

> *"I think perhaps one thing that could be done is that the team leader says you've marked so many, I want to see this… so the assistant examiner doesn't present their fifty, but could be called on to send any"*.

Whilst theoretically this may be preferable, the practical difficulties of adopting such a system were acknowledged. For example, the same examiner commented, *"I would imagine all exam boards are faced with the fact that this stuff has to be done in a reasonably short space of time, and that is a problem"*.

To summarise, parents, students and teachers appear largely unaware of current procedures for ensuring that quality of marking is maintained, suggesting a lack of knowledge and understanding of marking procedures. All these participants supported a system of monitoring and checking the marking of examiners though, possibly by using a system of regular checking. How this type of system would deal with discrepancies in marking is an important issue however, given that it could potentially discover a large number of discrepancies between senior and assistant examiners. Presumably, a larger sample from that particular examiner would need to be re-marked to determine how accurately the examiner is marking. The examiners tended to support the current script sampling procedures, believing it is sufficient to ensure quality of marking. The recognition of the practical constraints on the examination system meant examiners struggled to suggest alternative ways in which quality of marking could be assured.

The re-marking of examination scripts: a confirmatory vs. independent look
Although procedures are in place to ensure quality of marking is maintained, situations arise when centres or candidates are dissatisfied with the mark received and therefore request that the script is re-marked. According to the QCA Code of Practice (2007, pg 67) a re-mark is defined as *"a process by which a second examiner reviews the marking of the first examiner to make sure that the authorised mark scheme has been applied reliably"*. The examiner undertaking the re-marking is usually more senior than the original examiner, and is able to view the marks and comments of the original examiner during the re-mark. When undertaking a re-mark examiners are essentially instructed to check the reliability of the marking of the original examiner; if they believe the mark originally awarded fairly reflects the script's merit then it should remain the same. If this is not the case, a mark may be raised or lowered.

On completion of a re-mark the centre/candidate is informed of the resultant mark/grade. In situations where they not satisfied with the decision, three stages of appeals are possible, culminating in an appeal with the Examinations Appeals Boards (EAB) who operate independently of the awarding bodies and QCA. The task undertaken by the EAB is to determine whether the awarding body in question used the correct procedures throughout the marking process, rather than actually re-mark the script itself. The following section describes the findings from discussions with parents, teachers and students concerning perceptions of the procedures adopted by awarding bodies for re-marking examination scripts, as well as opinions on how re-marking should be undertaken. Satisfaction with the current re-marking system, and suggestions on how the system might be improved was discussed with examiners.

Parents, teachers and students involved in this research were all aware re-marking of examination scripts was possible. This might suggest an improvement since 2000, when Edexcel (2000) found only 76 *per cent* of parents were aware re-marking was possible. Notably however the research undertaken by Edexcel included the views of a much greater number of individuals than the present study. Despite awareness that re-marking is possible, perceptions of the actual procedures used for re-marking among parents, teachers and students varied considerably. Respondents rightly tended to assume the paper was re-marked by someone other than the original examiner, who was more senior. For example, one teacher responded, *"I would have thought it would have gone back to a more senior examiner to be marked"* and another *"I would imagine it went to someone more senior than the initial paper went to"*. Perceptions of what actually happens when a script is re-marked, including how many examiners would re-mark a script and how the final mark would be assigned differed though. The opinions of the students ranged from those who believed one examiner would re-mark each script, to those who thought four or five examiners would re-mark each script, where the final

mark constituted an average of each examiner's mark. Similarly, the perceptions of parents varied considerably. One parent believed just one examiner re-marks each script and responded, *"I would assume it would just be one person"* whilst others assume more than one examiner would remark each script. For example, one parent commented, *"I would imagine it perhaps was double checked"* and another, *"I would think there are people checking the re-markers"*. Compared to students and parents the perceptions of teachers were much more in line with current procedures. They tended to assume just one more senior examiner would re-mark each script.

Discussion also focused on whether parents, teachers and students think re-marking is a confirmatory approach, where the examiner re-marking the script can see the marks and annotations of the previous examiner, or an independent look at the work. Overall most of the participants thought re-marking involves an independent look at the work, in the absence of any previous marks and comments. For example, one parent commented, *"I think it would be a completely independent look…because I think you would be influenced otherwise"*. Most candidates tended to believe re-marking is an independent approach, arguing that being able to see the original mark may influence the examiner undertaking the re-marking. The idea of adopting an independent look at the work would also be the preferred method of re-marking for teachers, parents and candidates. For example, one teacher responded, *"A complete from scratch mark with no influence by the previous marking given…that would be an ideal system"* and another, *"It should be an independent look"*. However, these participants tended not to take account of the practical difficulties posed by adopting an independent approach to re-marking scripts.

Satisfaction with current re-marking procedures was also discussed with examiners. As those who took part were assistant examiners or team leaders, none had any previous experience of re-marking. Therefore, this is a topic most had not previously given much consideration. For example, one examiner commented, *"I don't know whether there's a problem…I haven't thought about re-marking as such"* and another *"I hadn't really thought about it"*. Examiners tended to agree however that a confirmatory approach for re-marking examination papers is satisfactory. One examiner in support of a confirmatory approach argued, *"I do think it's probably a reliable system, because the original [examiner] may have missed something".* Another suggested,

> *"The task in hand is that the student is appealing against the mark given so you need to see what that mark is…we are urged to put comments on quite a bit and you need to see the reason why that mark has been given. I don't think going back to the naked script would be very useful".*

In summary, although parents, teachers and students are aware re-marking is possible, most were unaware of the actual procedures involved. Most participants supported an independent approach to re-marking, although failed to acknowledge the practical difficulties associated with this. This may stem from these individuals being unaware of how a re-mark is currently defined by QCA, which essentially informs the procedures used by awarding bodies. Examiners tended to support the current system adopted for re-marking scripts, probably due to their greater understanding of the purpose of re-marking.

Discussion

The aim of this research was to explore perceptions and opinions of marking procedures amongst key stakeholder groups in the examination system. Previous research into actual perceptions of marking procedures is notably lacking in the literature and those studies which have been conducted have only focused upon how informed individuals believe they are, and not individuals' knowledge of actual marking procedures themselves. Furthermore, past research has tended to focus on examination standards, and less so on marking procedures or awarding body operations. This research therefore aimed to extend the findings of previous studies, by offering an in-depth exploration of individuals' perceptions of marking procedures. Additionally, it also aimed to explore individuals' views of what marking procedures should be adopted by awarding bodies.

Overall, the findings of this research suggest generally there is a lack of knowledge of marking procedures amongst the individuals who took part. This suggests either there is a lack of information in the public domain regarding awarding body operations, or this information does exist but is not sufficiently visible or accessible to the general public. Alongside this lack of knowledge of marking procedures there also appears to be a lack of understanding of the examination system, resulting in unrealistic expectations of how examination papers should be marked. Many of the parents and students who took part in this research tended to expect that examination scripts are subject to either double or even triple marking. Although perhaps a theoretical ideal, these procedures would be practically near impossible to adopt. There are shortages of examiners even under the current system and the financial implications of having more than one examiner mark each script would be enormous, factors many participants in this research understandably failed to consider. This implies a lack of understanding of the scale of the examination system, possibly resulting in false expectations and misconceptions about the system and a lack of confidence and trust in awarding bodies.

These findings suggest awarding bodies are at present not transparent enough with regard to their operational procedures. Arguably, increased transparency by awarding bodies would result in increased understanding and ultimately confidence in the examination system. For example, findings by QCA (2006a, 2006b) suggest increased understanding of the examination system is associated with increased confidence in that system. Despite this, the robustness of the evidence supporting this claim is perhaps somewhat questionable, therefore highlighting a need for further evidence detailing a link between greater understanding of, and greater confidence in, the examination system.

It has been suggested by some individuals that increasing understanding and transparency of the examination system, including issues relevant to measurement inaccuracies may adversely affect public confidence in the examination system. Newton (2005) however disagrees, claiming that '*not understanding errors in measurement error offers a greater threat than understanding it*'. He argues there has been a lack of information available regarding the strengths and weaknesses of the examination system, causing unrealistic expectations among the general public which are not met. The findings from the present study appear to support this claim, showing that individuals lacking an understanding of the scale of the examination system tend to adopt expectations of marking procedures that are not feasible. In line with the views of Newton, this study therefore supports the argument for greater education of the general public and greater transparency of awarding body operations. Not providing sufficient information to stakeholders in the examination system can lead to misunderstandings and possibly a lack of confidence and trust in the system.

Newton (2005) offers several suggestions of how to increase transparency of the examination system and gain a greater public understanding of measurement inaccuracy. Firstly, he suggests more research should be conducted into the reliability and validity of examinations in the UK, in such a way that the findings can be communicated effectively to key stakeholders. Secondly, the findings from such studies should be published in the form of defensibility arguments, written for, and accessible to, the general public. Thirdly, statements on the strengths and weaknesses of assessment should be published, in an attempt to increase the publics' understanding of assessment.

The lack of knowledge about awarding body operations amongst stakeholders in this study exists despite emphasis for improved transparency of awarding body procedures following the A-level results crisis in 2002. Recommendations included a need for much clearer communication on how exams are marked and graded, as well as a need to ensure teachers, students and parents have the right levels of information about awarding body procedures. Findings from this research suggest these recommendations are not being fully met. Similar to previous studies (Edexcel, 2000; QCA, 2006a, 2006b), many stakeholders still do not seem to receive information about awarding body procedures. This is despite QCA attempting to raise awareness of awarding body procedures by developing an official guide to A-levels (QCA, 2005a) and GCSEs (QCA, 2005b).

As well as increasing the availability of information to the general public, awarding bodies could also benefit from becoming more visible in the public domain. At present the publics' attention only tends to be drawn to awarding bodies by the media. The national press readily highlight any errors or problems encountered during each examination series and every summer the public are exposed to media stories of falling examination standards. These stories rarely, if ever, shed a favourable light on awarding bodies, so can only serve to reduce public confidence in the examination system. In the present study for example, parents recalled media stories of errors and concerns associated with the marking of examination scripts, where as none recalled any information received from the awarding bodies themselves. Compared to the media the awarding bodies themselves do less to promote their activities. This however, within AQA at least, has begun to change more recently. Key figures within the organisation have more actively engaged with the media around the examination period, therefore giving the media something to report and refer to. If awarding bodies do lack visibility in the public domain this can leave them exposed to media attacks during the examination series, which invariably undermines confidence in the examination system.

While this study gives an in-depth exploration of perceptions of awarding body operations among key stakeholders, there are three limitations. Firstly, the research necessarily focused on a small number of individuals, meaning that the results are not equally generalisable across and within each stakeholder group. Secondly, the complexity of the issues involved and the fact that many participants had little understanding of the examination system meant discussions were sometimes problematic. In some situations prompts had to be used to elicit a reply from participants which may have influenced an individual's response, although attempts were made to avoid this. Finally, considering the interviewer also transcribed the results, it is possible their own opinions of what seemed important influenced the reported findings. Despite these limitations, compared to previous research this study gives a much more in-depth discussion of the issues surrounding the marking of examination papers. The use of semi-structured interviews and a focus group meant respondent's ideas could be explored fully, and interesting responses could be discussed in more detail.

Conclusions and recommendations

This research suggests key stakeholders are largely unaware of the operations of awarding bodies, including the procedures adopted for the marking of examination scripts. Although QCA have produced guides to the GCSE and A-level system, it is clear from this research as well as past studies that only a proportion of teachers, parents and students receive these. Thus, it is recommended that this information needs to be much more visible to stakeholders, as well as the general public and that awarding bodies should also seek to be more transparent about their routine processes. A lack of transparency in the system can result in suspicions and mistrust in awarding bodies, which can be damaging to an awarding bodies reputation and therefore their overall business.

The findings of this research also suggest that there is a lack of understanding of the examination system among stakeholders. Thus, increasing information available on marking procedures would be best accompanied by explanations of why the current system of marking is adopted, and why having extended double, or multiple marking would in practical terms be unfeasible (Fearnley, 2005). It would appear simply informing individuals of marking procedures is not sufficient; it is also important to educate individuals and increase their understanding of the system.

Although it has been suggested increased understanding of the examination system would result in greater confidence in the system, the evidence upon which this notion is based has been questioned. Therefore, perhaps what is needed first is evidence supporting this link between greater understanding and greater confidence and trust. Once this is established, awarding bodies might be better equipped to begin increasing awareness and understanding among key stakeholders.

These recommendations raise the issues of how awarding bodies should together communicate this type of information most effectively to centres, students, parents and the general public. Ensuring the information is presented in the right way and is accessible enough is of utmost importance if awareness is to be raised successfully. Equally, determining the right levels of information to provide to individuals is important. Giving individuals too much information may cause confusion, whilst not providing enough information may still leave individuals with uncertainties about the system. Finally, ensuring any information that is conveyed to stakeholders and the general public is understood by the recipients as intended is vital for increasing understanding, and ultimately confidence in the examination system.

Rachel Taylor
May 2007

References

Baker, E., McGaw, B. & Lord Sutherland of Houndwood (2002) *Maintaining GCE A Level standards: The findings of an independent panel of experts*. London: QCA

Edexcel (2000) *Parents' perceptions of GCSE & A-level exams in England & Wales*. London: Edexcel

Fearnley, A. (2005) An Investigation of targeted double marking for GCSE and GCE, A report based on a re-marking of GCSE English and GCE Business Studies Summer 2004 scripts, organised by the AQA as part of an NAA series of Investigations into the Enhanced Quality of Marking.

Meadows, M.L. and Baird, J. (2006) What is the right mark? Respecting other examiners' views in a community of practice. AQA Research Committee paper, 14 June 2006, RPA_06_MM_RP_019

Newton, P.E. (2005) The public understanding of measurement inaccuracy. *British Educational Research Journal,* v31, n4, pp 419-442

QCA (2005a) *AS and A levels – the official student guide to the system*. London: QCA

QCA (2005b) *GCSEs – the official student guide to the system.* London: QCA

QCA (2006a) *GCSEs and A levels: the experiences of teachers, students, parents and the general public*. Research study conducted for QCA

QCA (2006b) *GCSEs and A levels: the experiences of teachers, students, parents and the general public*. Research study conducted for QCA

QCA (2007) *GCSE, GCE, GNVQ and AEA Code of practice*. London: QCA

Response to the Tomlinson report by an AQA spokesperson (3 Dec 2002) Retrieved 23rd April 2007 from **http://www.epolitix.com/EN/ForumBriefs/200212/1B584136-D860-48EB-B3B8-2BEF8E1B6325.htm**

Spearman, C. E. (1904a) 'General intelligence' objectively determined and measured. *American Journal of Psychology*, v5, pp 201-293.

Spearman, C. E. (1904b) Proof and measurement of association between two things. *American Journal of Psychology*, v15, pp 72-101.

Tomlinson, M. (2002) *Inquiry into A-level standards. Final Report*

APPENDIX A SUMMARY OF MAIN TOPICS DISCUSSED BY PARTICIPANT GROUP

How are exams marked: the hierarchical vs. consensus approach to assigning the true mark

Students/parents
- What happens to examination scripts once the examination has finished?
- Who marks these examination scripts/who are examiners?
- What happens when an examination script is marked? (How many examiners mark each script? How is a candidate's mark assigned?)
- If this was you/your child, how would you like your/their mark to be assigned?

Teachers
- How do you think your student's mark on an examination paper is assigned?
- How would you like your student's mark on an examination paper to be assigned?
- Do you think theoretically the hierarchical approach[1] adopted by UK awarding bodies is a satisfactory approach to determining a candidate's 'true' mark?
- Assuming unlimited resources, would you rather see a consensus approach[1] to determining a candidate's 'true' mark?

Examiners
- Putting aside any practical issues, do you think theoretically the hierarchical approach[1] is the best way to estimate a candidate's 'true' mark?
- Assuming unlimited resources would you prefer a consensus approach[1] to determining a candidate's 'true' mark?
- Benefits/Limitations of the hierarchical/consensus approaches

The reliability of marking

Students/parents/teachers/examiners
- Imagine a situation where two examiners marked the same examination script, would you expect them to give the same mark?
- Do you think there is any room at all for disagreement between examiners?

How is quality of marking assured?

Students/parents/teachers
- Imagine a candidate's script was sent to examiner X to be marked. Do you think the quality of this examiner's marking is checked throughout the marking process?
- Who do you think checks the examiner's marking?
- How is the quality of marking checked?
- How would you like quality of marking to be assured?

---

[1] The hierarchical and consensus approaches to determining a candidate's 'true' score were described and explained to teachers and examiners.

- What do you think happens if an examiner is found to be marking out of line with the required standard?

Examiners
- Do you think the current script sampling processes are sufficient to ensure quality of marking is maintained?
- How would you like quality of marking to be assured?

The re-marking of examination scripts: a confirmatory vs. independent look

Students/parents/teachers
- Imagine our candidate did not receive the mark they anticipated on their examination; do you think there is anything they can do in this situation?
- What would you do if this was you/your child/your student?
- What do you think happens when a piece of work is re-marked? (Who does the re-marking? How many examiners? How is the final mark assigned?)
- Do you think the person re-marking the script can see the marks and comments of the original examiner?
- What would you like to see happen when a piece of work is re-marked?
- Imagine our candidate was still not satisfied with the mark received following the re-mark, do you think there is anything they can do now?

Examiners
- Do you think current procedures adopted for re-marking examination scripts are satisfactory?
- Do you think re-marking should be a confirmatory approach, or do you think it should involve an independent look at the work?
- What procedure for re-marking scripts would you find most satisfactory?