Ofqual Quality of Marking Qualitative Research Study

Final Report

February 2014

Contents

1. Executive Summary

In April 2013 Ofqual commissioned Oxygen Brand Consulting Ltd to conduct a qualitative research study with examiners, teachers and head teachers as part of its review of Quality of Marking in general qualifications[1]. The aim of the study was to understand the attitudes of teachers and examiners to the marking process, where they felt it could be improved and how more teachers could be persuaded to become examiners.

The fieldwork, which took place during May, June and July 2013, consisted of four group discussions and 46 interviews. In total 82 individuals were surveyed: 32 teachers or head teachers who were not also examiners, and 50 examiners representing the main exam boards in England[2], of whom 32 were current teachers.

Teacher findings

Teachers believed the overall quality of marking of external exams in England is high. They perceived only a small level of error and believed this was largely accounted for by the fact that the marking of some subjects is by nature subjective.

The teachers most likely to disagree that marking was generally accurate were:

- Those in schools[3] (mostly selective state and independent schools) with the greatest pressure for students to perform at the highest level.

- Teachers of subjects such as English, history, psychology and sociology, where papers include questions requiring extended responses (worth 30-40 marks) and where a high level of individual examiner judgment is necessary.

- Teachers affected by lower grades than they expected in the 2012 GCSE English results.

---

[1] These include GCSEs, IGCSEs, A levels, the International Baccalaureate (IB) Diploma, Pre-U Diploma and International A levels.

[2] Assessment and Qualifications Alliance (AQA), Cambridge International Examinations (CIE), the International Baccalaureate (IB), Oxford, Cambridge and RSA Examinations (OCR), Pearson Edexcel and Welsh Joint Education Committee (WJEC).

[3] For brevity, throughout this report we will use the term "schools" to refer to schools, colleges and all other institutions where students study general qualifications.

Teachers in non-selective maintained schools and less experienced teachers are more likely to hold positive views about the quality of marking than more experienced teachers and teachers in selective state and independent schools.

Teachers in non-selective maintained schools were most concerned with the quality of marking as it affected their whole cohort, particularly at the C/D grade boundary at GCSE. In contrast, teachers in selective state and independent schools tended to be more focused upon individual student results, particularly at grades A*, A and B at A level.

Teachers told us that they mainly judge the quality of external examination marking by comparing students' achieved grades and marks to those that the teacher predicted for them. They also base their view on recalled scripts.

Most teachers in our sample had little or no understanding of how the technicalities and checks of the external marking process work. In particular teachers do not know about the ways in which exam boards monitor the quality of marking carried out on-screen. Our research suggests teachers would have more confidence in the quality of marking if they had access to more information about the marking process.

There was a perception amongst both teachers who did and did not examine that examining gave teachers access to additional information that helped them to prepare their students for exams.

Teachers believed the chief determinant of quality of marking was good examiners and therefore wanted most examiners to be teachers teaching the specification that they were examining.

Examiner findings

Overall, examiners[4] in this study believed that external exam marking is accurate. Examiners tended to describe themselves and other examiners as conscientious individuals, doing a job which they believe carries profound responsibility.

Senior examiners observed that the system of marking in England is in radical transition, moving from face to face standardisation of examiners and pen and paper marking (systems that had been in place for decades) to online standardisation and on-screen marking.

There was strong concern among examiners about the increasing adoption of online standardisation, particularly amongst examiners of papers including questions

---

[4] For an explanation of the terms used to describe different types of examiner, see page 8.

requiring extended responses from students and subjective judgments from examiners. The main criticisms of online standardisation were:

- It does not give examiners sufficient opportunities to ask questions and clear up any misunderstandings about the application of the mark scheme.

- Face to face standardisation encourages discussion of a large number of possible alternative interpretations of the question. As a result, when the examiner finally came to mark they would have discussed a large number of the potential answers and would be able to deal with them.

- It removes a key source of feedback for senior examiners. Meeting with current teachers enables the senior teams who write the papers (who are less likely to be current teachers) to keep in touch with the classroom.

- It may have an adverse effect upon the retention of examiners. Senior examiners felt online standardisation is less rewarding and instructive for examiners as well as more isolating.

Examiners were generally supportive of on-screen marking and felt it improves accuracy of marking for exams with short questions with few marks or unambiguous answers. However examiners thought that, for longer responses, the current software makes it too hard to get an overview of an essay and annotate it correctly.

The main suggestions from examiners for improving the quality of marking were:

- Return to face to face standardisation for some paper types. If that is not possible, face to face training should be provided for new assistant examiners.

- Review the team leader role to address the risk of exam boards becoming over-dependent on team leaders to supplement online standardisation with detailed personal guidance. Also introduce consistent guidelines on how team leaders should communicate with and support assistant examiners.

- Improve the recruitment, retention and development of examiners.

- Give examiners more feedback on their performance at the end of the marking process.

How to attract more teachers to the examining profession

On the whole the teachers in our sample had a negative perception of examining as a stressful and unpleasant experience. This view was reinforced by feedback from other teachers who had not enjoyed examining, and by a perception of a shortage of examiners, giving the impression that 'nobody wants to examine'.

The attitude of a teacher's school to examining has a large effect on that teacher's desire to become an examiner. Some head teachers (mostly, but not entirely, in this sample, in independent and selective state schools) have a positive attitude to examining and their schools adopt a highly strategic approach to using examining as a route to improved exam results. Other head teachers have a negative or passive attitude to examining which discourages their teachers from becoming examiners.

Teachers felt that exam boards should emphasise the possible benefits of examining to a school's exam results. This was seen by teachers as the main incentive for a teacher to become an examiner. Teachers also felt that exam boards should do more direct outreach activities with schools, local authorities and cluster groups to get their message across to teachers and head teachers.

2. Background

In April 2013 Ofqual commissioned Oxygen Brand Consulting Ltd to conduct a qualitative research study with examiners, teachers and head teachers as part of Ofqual's Quality of Marking review. The objective of this research was to gain a deeper understanding of the motivations, beliefs and attitudes that underpin teachers' and examiners' perceptions of quality of marking.

The study covered the themes below.

Amongst teachers and head teachers:
- Perceptions of the quality of marking of general qualifications in England
- Levels of understanding of the marking process
- Information that teachers would like to receive from exam boards
- How marking of exams might be improved
- How teachers feel about the role of examiner
- How to attract more teachers into examining

Amongst examiners:
- Perceptions of the quality of marking of general qualifications in England
- Perceptions of the marking process
- Difficulties and challenges that examiners experience in marking
- How marking of exams might be improved
- How to attract more examiners (including teachers) into examining

3. Sampling and Methodology

Oxygen designed a research structure consisting of both group discussions and depth interviews.

1. **Group discussions** to help stimulate debate and discussion, allowing peers to discuss the themes amongst themselves and prompting greater depth of thought as new examples and perspectives were suggested by respondents. 36 individuals took part in these discussions (four sessions each of nine participants).

2. **Depth interviews** to deliver greater geographic coverage and variation in types of examiner and teacher. The interviews allowed us to collect the views of stakeholders who would struggle to attend a centrally located session, who were teaching harder to find qualifications, or whose presence in a group might distract other people present, for example, because of their relatively high professional status. 46 individuals took part in these interviews, which were mostly carried out by telephone.

In total 82 individuals were surveyed between May and August 2013. 32 were teachers or head teachers[5] who were not also examiners, and 50 were examiners, of whom 29 were current teachers, and most of the remainder were retired teachers.

The participants of this study reflect a range of different types of school, levels of examiner and teacher seniority, and exam boards providing general qualifications in England. However, these individuals represent a very small proportion of the 51,000 examiners who mark exam scripts each year and the more than 500,000 teachers currently working in schools[6]. Therefore some of the observations reported in this study may not be representative of the views or experiences of the wider body of teachers and examiners.

A further 17 depth interviews with examiners were conducted later in the study, to explore emerging findings from Ofqual's quantitative survey[7] that indicated examiner dissatisfaction with the decline of face to face standardisation. These interviews

---

[5] For brevity, teachers and head teachers will be referenced as 'teachers' in the remainder of this report, except in any cases where head teachers' views differ from the views of teachers.

[6] Department for Education (November 2012) "School Workforce Census", total number of teachers (FTE) working in publicly funded schools in England. Independent Schools Council (January 2013) "ISC Census 2013", teachers (FTE) employed at ISC schools in the UK.

[7] Ofqual (2013) "Review of Quality of Marking in Exams in A levels, GCSEs and Other Academic Qualifications. Findings from survey of examiners, May 2013."

explored examiners' perceptions of how reduced use of face-to-face standardisation may affect the quality of marking.

Across both methods, the participants were mainly teachers and examiners of the subjects and exam types associated with the highest levels of dissatisfaction with marking. This helped to ensure the study provided as much insight as possible into potential improvements to marking. This means the findings of this study focus on particular problem areas and are not representative of the views of teachers and examiners in general, and may not agree with the findings of the quantitative research that Ofqual has conducted in parallel with this study. However our sample did include a smaller number of teachers and examiners of subjects and exam types with much higher levels of satisfaction with marking, in order to provide a point of comparison.

Further details of our methodology are provided in the appendices to this report.

In this report we use the term 'examiners' to refer to all of the examiners, at all levels of seniority, who took part in this study. We use specific terms to describe the following sub-groups of examiners:

- 'Senior examiners' are Chairs of Examiners, Chief Examiners and Principal Examiners.

- 'Team leaders' are examiners with responsibility for managing a group of assistant examiners.

- 'Assistant examiners' are those examiners without a supervisory role. These examiners are often simply called 'examiners' but some exam boards use the title assistant examiner. We use this term in our report to make the distinction clear between these examiners without supervisory responsibilities, and the wider group of examiners at all levels of seniority.

Presenting perceptual data

This report presents a discussion of the perceptions of teachers and examiners of the marking of external exams. These perceptions cannot be assumed to give an accurate reflection of the marking system in England. Throughout our interviews, both teachers and examines revealed varying degrees of understanding of the marking system. Nonetheless, every response has been treated equally and the data has been presented as it is. For a discussion of how these perceptions correlate with other sources of evidence, further reading of Ofqual's *Review of Quality of Marking in Exams in A levels, GCSEs and other Academic Qualifications Final Report* is recommended.

4. Teacher findings

This section of the report summarises the findings from discussions and interviews with teachers. It focuses mainly upon the views of the 32 teachers surveyed who do not work as examiners, and also references contributions from the 29 individuals who work as teachers and examiners.

Teacher perceptions of quality of marking of academic qualifications

Teachers defined high quality external marking as:

■ Marking leading to final grades that match those they have predicted for their students.

■ Marks that reflect teachers' understanding of their exam board's published specifications and mark schemes.

■ Marks (and resultant grades) that are, in the words of teachers, 'consistent' or 'predictable.' This judgment is based on an approximate comparison of cohort ability from year-to-year, assuming a consistent standard of teaching.

Teachers perceived the marking of A level and GCSE exams, and exams in equivalent qualifications, to be 'good'. Several with knowledge of examining in other countries (teachers who had lived and/or taught abroad) drew on international comparisons to support this. The majority of teachers believed that about 95% of the marking of external exams is accurate, based on their predictions of what students ought to achieve. Their view was based on an overview of marking of all external exams in academic qualifications in their schools, not just in their own subject.

Most teachers saw the perceived level of error (around 5%) as almost inevitable in an operation of the scale of examining in England and which includes subjects of a 'subjective' nature. They were aware that these more subjective subjects have more issues with accurate marking than other subjects. Teachers therefore felt that the current level of accuracy is acceptable. Some expressed the view that to remove error completely would mean moving to much more structured questions which they thought was far too great a price to pay for complete accuracy as it would lead to a form of teaching and learning that they did not support.

Teachers also thought they observed regular (usually annual) 'blips' in marking that are spread around cohorts and which affect subjects and papers randomly.

Teachers who disagreed that marking was accurate mainly came from one of these three groups:

1. **Teachers in schools with the greatest pressure for students to perform at a high level.**

This group had the lowest levels of satisfaction but were nonetheless broadly satisfied. This group consisted of teachers, usually in selective state or independent schools, who reported they were expected to ensure that large numbers of their students attained the very top A level grades (A*/A/B) needed for entry to Oxbridge and the Russell Group universities. In these schools any perceived marking error receives detailed scrutiny and investigation. There is low tolerance even of small variations from the results they expect for their students and great efforts are made to secure year-on-year improvements in results, combined with 'predictable' results over time. These teachers report that their investigations with the exam boards are sometimes unsatisfactory or inconclusive and this adds to their frustration about perceived marking inaccuracies.

2. **Some teachers of subjective subjects and subjects with papers requiring extended responses to questions.**

We interviewed six teachers who said they had suffered serial issues with the marking of subjective subjects and subjects with papers requiring extended responses to questions (such as A level English literature or A level history). They were dissatisfied with the quality of marking within their exam boards or with the direction of marking within their subject across exam boards. However this belief usually related to the marking of their specific subjects and the marking of similarly subjective subjects within their school and not to all exams.

The other teachers in our sample who taught subjective subjects and subjects with papers requiring extended responses (approximately 14) were generally satisfied with the quality of marking but many said they had suffered occasional issues with the accuracy of marking.

3. **Teachers and head teachers affected by lower grades than expected in the 2012 GCSE English results.**

Around a third of teachers working in state schools interviewed referred (without prompting by our interviewers) to the issue of grading in GCSE English in 2012 as having affected their views of marking. They did not believe that the marking by examiners had necessarily been less accurate but they did believe, based on their assessment of individual students' outcomes, that the grading of GCSE English in 2012 had not been consistent across the year. Some teachers felt that this grading issue had undermined their confidence in the marking of GCSE English and also of other subjects.

Attitudes to quality of marking by school type

The research confirmed results from Ofqual's previous quantitative work that teachers in non-selective maintained schools and less experienced teachers are more likely to hold positive views about the quality of marking, whereas more experienced teachers and teachers in selective state and independent schools are more likely to hold negative views about the quality of marking[8].

Comprehensive schools, academies and free schools

Senior teachers and head teachers in these schools agreed they expected approximately 5% of exam results in any year to vary significantly from their expectations but believed the vast majority of exams are marked 'predictably' in line with previous experience and predicted grades.

In our sample, marking of GCSEs was believed to be accurate 'enough' with students getting 'roughly' what their teachers expected for them.

> *'It's there or thereabouts basically.'*
>
> Teacher, comprehensive school

> *'I think on the whole, marking is reliable. We don't have enormous numbers of marks that come back we think have been done unfairly. Every year we do re-submit and ask for* [re-marking] *and occasionally they do get re-graded.'*
>
> Head teacher, comprehensive school

Teachers said that their schools invested considerable resources in statistical analysis of their exam results to monitor 'macro' patterns and movements across school years and cohorts. It was noticeable that this was almost the sole focus in non-selective maintained schools. Compared to selective and independent schools, senior staff were much less interested in looking at results at the individual student level. If individual students diverged from expectation at GCSE or A level, teachers said that this sometimes wasn't investigated in detail, particularly if the student had achieved the overall grades they needed. Unexpected results might be put down to last minute effort, extra help received outside of school or 'wobbles', not external examining standards.

---

[8] Ipsos Mori (2013) "*Perceptions of A level, GCSEs and other qualifications - Wave 11. Perceptions of Teachers and the General Public.*"

The two comprehensive head teachers in the sample reported that marking was a more 'approximate science' than their counterparts in the selective and independent sector, believing that a degree of subjectivity in marking the more judgment led subjects was irreducible and perhaps even the price to be paid for a stimulating curriculum.

Teachers in non-selective maintained schools were most closely focussed on variations in marking at the C/D borderline at GCSE, particularly in the core 'English Baccalaureate' subjects[9]. This reflects the increasing importance of school accountability measures, which focus on the proportion of students achieving GCSE grades A*-C in core subjects. Schools report these results to their governing bodies and externally, and head teachers have targets to deliver on these results. In addition, many teachers regarded a C grade as a 'pass' and believed that students' life chances were greatly enhanced if they held a C grade rather than a D.

> *'A 'D' is a disaster for them… If they are not way off we appeal* [enquiry about results].*'*
>
> Teacher, comprehensive school

All head teachers in this sample were also fully aware of the statistical likelihood, based on the use of marking tolerances, that marks could increase as a result of a review of the marking (an Enquiry about Results)[10]. They saw some such enquiries as 'a one way bet' therefore.

Teachers in non-selective maintained schools also observed unpredictability in results among students aiming for top A level grades but said they were less likely to investigate these anomalies because their limited resources were more focused on the C/D boundary at GCSE.

Selective state schools and independent schools

Selective state schools and independent schools in this sample tended to focus on the 'end game' of Russell Group or Oxbridge admissions and therefore needed to

---

[9] A performance measure which recognises whether students have achieved a C grade or above across a core of subjects (English, maths, history or geography, the sciences and a language. See: http://www.education.gov.uk/schools/teachingandlearning/qualifications/englishbac/a0075975/the-english-baccalaureate.

[10] If a school is concerned about the accuracy of a student's result, it can ask the exam board to investigate the marking. This is carried out through the Enquiry about Results process.

ensure many of their students achieved the highest grades at A level: A*/A/B (or equivalent in IB or Pre-U Diplomas). They were therefore more aware of and more intolerant of any variation in marking than non-selective maintained schools.

Teachers in independent schools in particular reported they were under enormous pressure from fee paying parents to ensure students secured admission to the 'very best' universities. These schools have a simple business imperative to be precise in exam preparation and investigation for each individual student.

This puts pressure on teachers and some told us that they approach the exam season with trepidation, cumulative frustration and a strong awareness of past 'incidents'.

> *'I often enter the exam season very fretful the day before* [results are issued] *that I am going to embark on an appeal* [enquiry about results]*…'*
> Teacher, independent school
>
> *'Every year, something. You don't know where it is going to go wrong. All you know is that it will somewhere, it is never one board with all subjects, it is one subject with one board, with one aspect of the paper. But all hell will break loose…then ideally you need to be better armed by the boards to talk to parents about* [the exam boards'] *processes.'*
> Head teacher, selective state school

Although all types of schools actively sought out qualifications they thought would deliver more 'predictable' results, selective state schools and independent schools in our sample did so with a greater frequency than maintained schools. Some independent schools particularly sought new qualifications they hoped would deliver 'better quality of marking' particularly at the top end of the grade range where they wanted their students' abilities to be recognised consistently. Amongst independent schools and selective state schools, head teachers' and senior teachers' opinions differed greatly on the reliability of the domestic exam system. Some felt very negative, for example, about the accuracy of mainstream domestic exam boards compared to the International Baccalaureate (IB) qualification (which enthusiasts believed focussed greater resources on examining than the mainstream exam boards). Other head teachers and senior teachers felt marking in the mainstream system in England (GCSEs and A levels) was satisfactory, as well as being fair and straightforward to enquire into if they suspected a marking inaccuracy.

> *'It* [A level] *is accurate, but could be made even more so.'*
> Teacher, independent school

> *'I am not one of those head teachers who thinks the whole system is broken.'*
>
> Head teacher, independent school

Drivers of teacher perceptions of quality of marking

All of the teachers in our sample told us that they judge quality of external marking by comparing students' actual results to their predicted grades and marks. Teachers told us that they invest a great deal of time and effort in attempting to understand published mark schemes and specifications so they can be as accurate as possible in their own marking of students' work and provide guidance to students on how to improve their results. Some schools even invest in sending teachers to training sessions run by the exam boards. Teachers apply this information to their marking of students' work throughout the year, not just at the time of predicting grades.

The predicted grades resulting from teachers' internal marking are monitored by senior teachers and shared with students and parents. Teachers said they find it difficult to understand the language of published mark schemes and specifications, but work hard to overcome this barrier to ensure that their predictions are based on exam board guidance.

Most teachers also said they 'feel they know' what students should achieve based on their more qualitative classroom observations over the course of a year.

To contextualise this, schools' predicted grades were sometimes felt to be 'optimistic' by senior examiners. This view was particularly held about small schools where senior examiners believed the school sees too few students to accurately judge the national standard[11].

Several teachers described a number of other experiences that had directly reduced their confidence in the accuracy of marking:

---

[11] Ofqual has requested that we note that there are a number of reasons, unrelated to marking, why students may not achieve the grades predicted for them by their teachers. For example, teacher optimism or pessimism when preparing predictions, or particularly effective or ineffective revision by students. It is also noted by Ofqual that the 'predictability' of a student's final grade (the extent to which a teacher has confidence that a student's final grade will match his/her predicted grade) is different from the predictability of the exam itself, which is the extent to which the content of the exam paper can be accurately predicted by teachers. This is a feature of the exam paper and not of exam marking.

- The mark changes substantially following a review of the marking as part of an enquiry about results. This, in teachers' view, 'proves' that errors can occur.

- The teacher reviews the exam script and deems it worthy of a higher mark than it received, even after a re-mark[12].

- The teacher reviews the exam script and believes the marking is not in line with the mark scheme and written guidance (that they have scrutinised in detail).

- The teacher believes an exam script has been over-marked (although schools do not complain to the exam board when they suspect an instance of this or believe they see it on a returned script).

- Repeated experience of re-marks, either individual or cohort, that do not result in a change to the final mark. This can reinforce dissatisfaction with the marking system, inculcating a belief that 'the system' is predisposed to uphold the original mark, rather than admitting the original marking was inaccurate.

Several teachers also described a number of influences that indirectly undermined their confidence in the accuracy of marking:

- Lack of understanding of the multiple checks and balances that exist to monitor the accuracy of individual examiners. Where an exam board strongly promoted its checking systems to teachers or where a school had a number of teachers who were examiners and knew the system, confidence tended to be higher.

- Common 'myths' about the quality of examiners, usually heard from other teachers (for example, 'exam boards use student teachers/undergraduates to mark A level scripts'). Some teachers also said they saw advertising for examiners, sometimes late in the exam cycle. This caused them to speculate that exam boards could be short of examiners and recruiting examiners with insufficient experience or who are not even teachers.

- Regular media coverage of 'poor marking' was said by a few teachers to affect their personal confidence in the quality of marking.

---

[12] The more limited annotation possible in online marking systems was also mentioned by a number of teachers who requested returned papers in essay based subjects such as English. When a teacher was dissatisfied with the mark, this further contributed to a lack of confidence as how the marks had been awarded was not seen to be sufficiently justified or transparent.

15

Teachers' understanding of the external marking process

Most teachers in our sample had little or no understanding of how the external marking and grading system worked in terms of quality control, checking or process. They could not describe what happened to a script once it left the school.

The knowledge or impressions they had were acquired from colleagues who marked now or had marked historically, from observing or talking to visiting moderators or from exam board training. Teachers interviewed from maintained schools said they had limited experience of exam board training sessions because of 'budget restrictions' or because it was harder for their schools to spare them to attend training.

Often teachers' understanding of external marking methods was outdated, meaning it was usually based entirely on pen and paper marking. Many teachers had no knowledge of the existence or increase of on-screen marking. Most thought their students' scripts were marked as a single batch by the same examiner and therefore looked for error patterns across the cohort, rather than at the level of individual students.

Those teachers who knew most about the system were:

- Teachers who were also examiners, (particularly senior examiners as they had some overview of paper setting, pre-standardisation and monitoring of assistant examiners).

- Teachers in schools with many examiners, but only if those examiners were sharing their knowledge proactively.

- Teachers whose exam boards invest in detailed communication about the marking process with schools. In this sample these teachers most commonly used the exam boards Cambridge International Examinations (CIE) and the International Baccalaureate (IB).

- Teachers who had switched exam board because of dissatisfaction with marking quality or predictability and who had researched how scripts were marked as part of that process.

On many occasions, those teachers who had the marking process explained to them had evidently retained detailed knowledge and it did appear to enhance their confidence in the exam boards they worked with.

Impact of improved knowledge on perceptions of marking

Teachers' confidence in the overall accuracy of marking almost always improved when information on the checks and balances in the marking system (see Appendix C) was shown to them at research sessions.

There was some scepticism among a small minority of teachers that every step of such a detailed process was carried out, but on the whole teachers felt the system of checking and monitoring was more sophisticated than they had assumed.

> *'It* [schematic of the marking process] *is a lot better than I thought. My mental image was that all the papers just went off in a pile to be marked by some old bloke in corduroys with leather patches on his elbows.'*
>
> Teacher, comprehensive school

There was a general consensus amongst the teachers in our sample that the most impressive elements of the process were spot-checking of assistant examiners' completed marking by senior examiners and the responsive and comprehensive monitoring of individual assistant examiners' performance that was possible with on-screen marking.

Two head teachers who said that they regularly deal with concerned parents at results time suggested similar information to that presented in the research sessions would be helpful, particularly if the proportion of scripts spot-checked or sampled in their school could be included:

> *'I would like at the time the results came out to see a report on that* [schematic of the marking process] *the day before. An email that said 'this is the process we went through, this is the national picture, and this is your school.''*
>
> Head teacher, selective state school

At the research sessions, teachers were also shown some statistics on the level of experience of examiners and on the number who were classroom teachers. They also found this reassuring as it addressed some myths and rumours about "unqualified" examiners and laid some of their latent fears to rest.

Information that teachers would like to receive from exam boards

Our research suggests that teachers would have more confidence in the quality of marking if they had access to more information about the process. However, when asked about the information they want from exam boards, teachers were not, unless

prompted, interested in information about the technicalities of the external marking process. They were much more focused on understanding how marks are awarded in exams so they can prepare students better, be more accurate in their marking of 'mock' exams and therefore make more accurate grade predictions.

Many teachers said they believe or 'know' that there is additional information or guidance only available to examiners that clarifies the mark scheme and what the language in it means. They said this is the reason why many schools encourage examining and why many teachers decide to examine.

Teachers believe this 'inside information' is hard for teachers to access if they do not examine:

> *'I said* [at a controlled assessment moderation visit] *'what are these things written down here on your sheet?', and the moderator said, these are the things your pupil has to do without which they cannot get the top band, these are the success criteria.'*
>
> Teacher, comprehensive school, drama

> *'This is it… the magic information you can only find out by examining.'*
> (Referring to a mark scheme annotated with guidance from a standardisation session)
>
> Teacher, comprehensive school, sociology

This view is almost universally supported by the perceptions of teacher-examiners, who (with the exception of some who marked only a restricted number of questions within a paper) all said that by examining they were able to guide students on how to answer questions in a way which would maximise marks, whether those students were higher or lower ability. These examiners did not believe that the same benefits could be derived from reading the specification or published example mark schemes. They claimed they had read the exam board literature repeatedly and in detail before they had started examining, but still suffered from misconceptions or confusion that was only resolved by examining.

Teachers of subjective subjects who were also examiners believed that they learned crucial extra information and techniques from examining which helped them with marking and teaching in school. They said they learned this best via face to face standardisation meetings or from their team leader briefing.

Most of the teacher-examiners in our sample also said they learned useful techniques indirectly by marking scripts from schools who in their opinion were highly skilled at helping students present answers to maximise marks.

Three teacher-examiners even claimed that a requirement in their exam board's specification was potentially misleading to teachers.

The small number of teacher-examiners marking only low tariff questions or 'items' in our sample said that they found this type of marking less useful for informing their teaching practice. These participants were also standardised online (see section five) and said that they found reviewing the written standardisation materials less useful in helping them to understand how to prepare their students for the exam than attending a face-to-face standardisation meeting.

Teachers who do not examine have access to everything published on their exam board's website including the published mark scheme, the level descriptors, the specification and the Examiner Reports, but not all feel they are able to understand them. Teachers in our sample said the material often employs 'jargon' and the guidance is sometimes 'vague' or contains insufficient examples to assist accurate internal marking. Some even said they suspected the guidance is deliberately vague. Three teachers said that when they telephoned exam boards with a query about marking they had been offered paid-for training rather than the explanation they wanted.

Teachers and examiners said that the Examiner Reports (published by exam boards for each unit in each exam series) describe how the mark scheme for each unit was interpreted against that year's questions and also explain typical responses and why they gained marks. The Examiner Reports were seen as valuable by the minority of teachers we spoke to who were confident that they understood them. However at least half of the teachers we spoke to did not know of their existence and of the others some had looked at them but found them hard to understand.

Teachers suggested some improvements that could be made to the materials that are already provided by exam boards:

- All teachers in our sample said there should be greater clarity of language in the published specification and mark scheme so that **anybody** reading it (including students) can understand it.

- One teacher said Examiner Reports should be released earlier so that teachers can incorporate any learning into their lesson planning for the autumn term.

Teachers also suggested some additional resources that they would find helpful:

- Many said they wanted access to the 'full guidance' on the mark scheme (that they believe examiners have access to) particularly concrete examples of the type of evidence or skills necessary to meet mark band definitions and precise explanations of language being used. One example given was the words used to define level descriptors, 'what is 'good' versus 'excellent' knowledge?'

> *'Is* [a student's work] *'insightful', A* or is it just 'exploring' that gets you an A – or…* [is it] *just 'understanding' that gets you a C?*
>
> Teacher, English literature
>
> *'What are "occasional mistakes", is one "occasional" or is it four?'*
>
> Teacher, music
>
> *'What is "thinking like a geographer" for the top mark band?'*
>
> Teacher, geography

- Many teachers said that they would find access to marked up past papers and small 'extracts' (from student scripts) helpful, to enable them to better understand the mark bands. Teachers said that as they work with student answers every day they would find this format helpful. Teachers also said this format would allow them to look beyond their own school at what 'the standard' is. Some teachers report Pearson Edexcel already adds extracts from essays to mark scheme guidance which they find useful.

- Access to free information from the exam boards, particularly for those teachers who could not get funding from their school for paid-for teachers' notes or training because they were teaching subjects that were felt by the school to be a low investment priority (for example GCSE drama, media studies).

- One teacher said more numerous examples of questions and model answers would be helpful when a new specification or question type is launched. When this point was raised in a group discussion, the other members of the group agreed and said they can feel 'at sea' with a new specification and feel students are vulnerable in this period. They said they wanted more examples to work through with students 'not just two or three'. Going through past papers and questions is common technique for exam preparation and teachers in this group said they wanted enough examples to cover both the initial explanation of the new requirement to students and then to help them practise before the exam.

Teachers noted that historically their departments had paid for scripts to come back from an exam board to keep as 'real' examples of a mark band, to enable teachers to better understand nuances of marking. However, approximately six teachers within

the sample (who taught subjects with papers that were marked on-screen) said they had recently asked for scripts to be returned and found them much less helpful because with on-screen marking there tended to be less annotation of scripts.

How teachers believe the marking of exams might be improved

Most teachers who were not also examiners found it hard to suggest specific improvements to the external marking process. This was understandable, as they had a limited understanding of the technical process.

As a result when prompted for suggestions for improvements what they tended to ask for was improvements in outcomes: making external marking even more 'predictable' (by this they meant consistency in results for 'similar cohorts' year to year), 'getting more of it right first time' and ensuring 'examiners are confident enough to award top marks and not just mark to the middle of the mark band'.

Because teachers felt they had received little technical detail from exam boards about marking and re marks, most simply assumed that errors occurred because exam boards employed 'rogue markers' and suggested that the boards identify and remove those examiners.

Teachers wanted all examiners to be current UK teachers of the specification they examine (and most assumed this was already the case), preferably experienced senior teachers. They wanted this in order to ensure an understanding of the specification by the examiner and what they saw as 'realistic' expectations of students' language and knowledge.

Several teachers, largely those with direct experience of examining, had some detailed suggestions about improving the marking of A level qualifications in subjective subjects and in questions requiring extended responses. These suggestions often focused on marking top performing students. However, there were also suggestions from a few teachers who had investigated specific incidents through recalling scripts, having detailed dialogue with exam boards, and submitting enquiries about results or appeals and who therefore based their thoughts on these investigations.

The chief concern of teachers is under-marking and harsh marking. Training was suggested to help such examiners understand concepts such as 'positive marking' or to disregard the influence of grammar or handwriting where this is not a valid aspect of the assessment.

Selective state and independent school teachers occasionally questioned the ability and/or knowledge of some examiners. Teachers of subjective subjects said examiners should be able to recognise a 'wider' or 'superior' knowledge from the student but were not always able to:

> *'A student who has written an A/A* answer who is so well informed the examiner can't recognise it.'*
>
> Head teacher, state selective school

An independent school head teacher, a former examiner, feared high performing students could be under-marked by examiners with 'degrees from less demanding universities' and claimed to have met teachers at standardisation they would not regard as suitable to teach in their own school. This is a factor behind some schools' moving to exam types and exam boards that they believe have a higher quality of examiner.

Six teachers said they believed or suspected examiners were sometimes unfamiliar with the set texts or optional topics that they marked:

- In English literature two teachers said they had seen indications from recalled papers that the examiner did not know the text discussed sufficiently well[13].

- Four teachers also felt multi topic subjects such as history, psychology and sociology could be vulnerable to inaccurate marking as a result of examiners' unfamiliarity with cases or topics.

Both assistant examiners and senior examiners corroborated that the current system *could* in their opinion, occasionally allow examiners to mark student responses relating to a set text or sociology or psychology topic with which they were not familiar. Three examiners in our sample said they had found themselves in this position in the past.

These teacher-examiners suggested that exam boards should set fewer topics and give marking of more rarely studied texts only to examiners who knew them well.

In our sample, all teachers of subjective subjects and subjects requiring extended responses to questions (such as A level religious studies and GCSE history) said their subjects had potential for original interpretations and arguments. They believed examiners occasionally do not cope with students who write a relevant but different or unexpected response to a question and said they wanted the examiner to reward good answers that met the requirements of the mark scheme but did not necessarily match the example answers provided in the mark scheme. Teachers feared

---

[13] One example given was from "An Inspector Calls", a standard text at GCSE, The candidate was aware that the play was first performed in 1945 though it is set in 1912. In his script he drew a comparison with the bombing of Dresden at a point when the characters praised aeroplanes. In the opinion of the teacher who had recalled the paper the examiner did not understand this. From their comment on the script this teacher believed the examiner was marking it as a text written in 1912.

examiners could be too dependent on specific examples in the exam guidance and lack the confidence to give high marks to unexpected answers.

We should note that whilst senior examiners in our sample agreed that this *could* happen, they explained that exam boards issue specific guidance about the necessity to award top marks when deserved, and assistant examiners are encouraged to escalate such queries to their team leader.

Team leaders said they guide assistant examiners to reject answers that are 'interesting' but do not match the mark scheme or demonstrate the skills or knowledge that the Principal Examiner intended to test when setting the question. However, team leaders also said they do educate assistant examiners to give full or good marks to unusual answers if they meet the mark scheme's requirements, for example, if the mark scheme says '…or any other relevant answer'.

5. Examiner findings

Overall perceptions of the quality of marking of academic qualifications

Overall, most examiners in this study believed that external exam marking is accurate. The newer and less senior examiners (assistant examiners) in the sample, however, regularly expressed doubts about their own performance. They hoped they were marking reliably, and certainly said that they try to, but are not always confident about this. They understand they are subject to surveillance by the team leader and the exam board but they do not always understand what is being done or when this surveillance takes place.

Senior examiners and team leaders had faith in the accuracy of the system:

> *'I feel the process rewards candidates appropriately. If I did not believe this I would not sign off the awards.'*
>
> Chair of Examiners, subjective subject (subject title withheld for reasons of confidentiality)

Senior examiners were particularly confident about marking accuracy because they have an overview of all the checks and balances in the system and are personally involved in setting papers, spot checking and supervising re-marking.

We came across only one senior examiner of the 17 interviewed who was dissatisfied with the current quality of marking on their paper. This individual felt online standardisation had compromised examiners' grasp of the specification and standard too far.

> *'I do have a problem with marking in my team. We no longer have face to face standardisation meetings and online standardisation does not work well. Examiners feel isolated, feel they ought to know and don't want to ask. There is not the personal relationship between marker and team leader that is important to engender trust, mutual respect and commitment.'*
>
> Chief Examiner, subjective subject (subject title withheld for reasons of confidentiality)

There was, however, concern among over half of the senior examiners about **future** quality and accuracy of marking, particularly centring on recruitment and retention of good examiners. They feared recent changes in standardisation methods were reducing the personal satisfaction and rewards examining had once provided.

A majority of senior examiners in subjective subjects felt the main exam boards were not recognising future risks adequately when they raised fears about them. These examiners also felt that exam boards were not taking sufficient action to accurately understand and investigate senior examiner concerns.

There was a consensus among senior examiners on what helped examiners mark accurately. Three factors were most felt to affect quality of marking:

1. **Effective standardisation**, defined as a shared and consistent understanding among everyone marking the paper of 'the standard':

    ▪ The Principal Examiner's intention in setting the questions, both in terms of the paper's overall philosophy and the specifics of the knowledge and understanding that students should demonstrate in their responses to the questions that have been set. Senior examiners said that in order to mark accurately assistant examiners must accept this intention rather than challenging it. Many called this 'psychological buy-in' to the mark scheme.

    ▪ A full and deep understanding of the mark scheme and how to apply it. A test of this was described as being able to deal confidently with original student answers or arguments not specifically mentioned in the guidance.

2. **Quality of examiners.** For senior examiners this meant not only recruiting the right individuals, preferably experienced current classroom teachers, but also retaining them and developing their skills over time. In the words of one senior examiner: 'a 'complete' examiner is not made in just a year.'

3. **Effective monitoring** of the quality of marking throughout the process. Quickly recognising, re-training or stopping inconsistent examiners who deviate from 'the standard'.

Examiner perceptions of the marking process

Standardisation

All assistant examiners and team leaders must be standardised before they are cleared to mark that year's paper. This means they have to complete training to ensure they understand the paper and mark scheme and then mark a number of scripts to an acceptable level of accuracy before they are cleared to mark.

There are three different methods of standardisation used by exam boards: face to face meetings, online 'self-standardisation' using mainly written materials, and online forums using voiceover and live online chat.

Some exam boards are increasingly using online standardisation in place of face to face standardisation. Many examiners in our sample expressed concern about this change and the effect it could have upon assistant examiners' ability to mark accurately and consistently.

Face to face standardisation

In this method examiners attend a meeting where the Principal Examiner introduces the paper and mark scheme and then examiners mark and discuss practice scripts in small groups with their team leaders. Assistant examiners and team leaders are then cleared to mark by marking a number of scripts in on-screen or hard copy format.

Examiners felt that the benefits of face-to-face standardisation over other methods were:

- It allows individual questioning by examiners of all aspects of the mark scheme and its interpretation, which examiners thought was extremely important for developing a full understanding of the mark scheme. Examiners said they felt being able to ask questions was the best way to ensure that they understood the mark scheme. 'A one way lecture on the paper' was, in the opinion of the majority, unlikely to be as effective as a session that allowed examiners to ask many questions.

- It allows all examining the paper to propose and discuss a wide range of possible interpretations of each question (and therefore to consider a wide range of possible answers that would be acceptable). Examiners felt that exposure to these different interpretations raised by other examiners, and to the judgment then delivered on those interpretations by the Principal Examiner or team leader, helped them to mark more confidently and consistently. This was seen as particularly valuable in subjective subjects where multiple interpretations are a key difficulty in marking accurately.

- It provides a fuller 'feedback loop' between schools and exam boards:

   o Assistant examiners (usually teachers) can feed back into both that year's mark scheme and into question setting and interpretation of the specification by the senior team in the longer term.

   o Senior examiners (many of whom in our sample were retired or not teaching the specification themselves) said they valued the opportunity to keep in touch with classroom practice for example the common textbooks used.

- It allows late amendments to the mark scheme. In practice examiners said amendments at this stage in the process are rare, but they do happen.

- It establishes a relationship between the team leader and 'their' assistant examiners, meaning the assistant examiner is more confident to call the team leader with queries during the marking process.

- Senior examiners and team leaders said it allowed them to identify assistant examiners who are 'not coping' or seem to not wish to adhere to the mark scheme or accept the intention of the paper.

- It offers additional training opportunities, such as special sessions for new assistant examiners before the main session starts. Examiners said they pick up useful tips on marking accurately and learn from observing the way experienced examiners order their work during sessions[14].

- It facilitates the development of future senior examiners within the system. Most senior examiners said they have responsibility for identifying new team leaders and said they use face to face sessions to identify individuals with the right interpersonal skills to make good team leaders and Principal Examiners in time. Senior examiners said that if they did not meet assistant examiners face to face then it was only possible to choose team leaders on the basis of excellent marking. Those who marked excellently might not have the people skills or nurturing attitude necessary to make good team leaders.

Most examiners, including many who were standardising online, expressed a firm preference for face to face standardisation on the grounds of ensuring a higher quality of marking. Leaving the issue of the inconvenience of attending face to face sessions to one side, this method was seen as the most effective at training examiners to understand the task of marking. It was also seen as by far the most rewarding from a professional development standpoint.

Online standardisation

In this method assistant examiners work through an online induction to familiarise themselves with the paper and the mark scheme. Most of the information is in written format although some exam boards also provide some video guides.

Certain types of examiners were happy with online standardisation. These tended to be those who examined relatively 'simple' papers with predominantly lower tariff

---

[14] For example, "it is useful to print out the final mark scheme and have it in front of you for the whole session to keep you marking consistently", and "re-mark the last script you marked the night before to check your judgment is consistent" (Pen and Paper method).

items, although online standardisation was not universally seen as satisfactory even among these examiners[15].

Other examiners satisfied with online standardisation included very experienced examiners marking a number of different papers and examiners who had been examining one paper for many years. They did not feel they needed any further training and were able to understand what was required of them from written instructions and conversations with their team leader.

Online standardisation also appealed to those examiners living a long way from major cities and those with small children. However, if those examiners were struggling or less experienced, they still said they would have been prepared to travel for face to face training.

However, online standardisation was the method most criticised by examiners of subjective subjects and subjects with papers requiring extended responses to questions. These examiners thought that assistant examiners' understanding of the mark scheme is potentially lower when they are standardised online, and that this can lead to inaccurate or over-cautious marking particularly on papers that include questions requiring an extended response. The main disadvantages of online standardisation were described as:

- It is 'lonely' and unrewarding. It limits the emotional and subject enhancement rewards that motivate a number of examiners. Senior examiners said they feel face to face standardisation has social, subject and professional development benefits that made the process of examining rewarding and retained good examiners in the system. Pay versus benefits in examining is a fine balance, and senior examiners fear online standardisation upsets the 'value equation'.

  > *'I think online standardisation contributes to the very isolation that a lot of* [smaller subject] *teachers are trying to counteract by deciding to mark and I think it will affect retention of new examiners. I think a lot of new examiners will go 'that was horrible' instead of thinking 'well, that was interesting as well as professionally developing it was also enjoyable.'*
  >
  > Chief Examiner, subjective subject (subject title withheld for reasons of confidentiality)

---

[15] Maths examiners said that "even in maths" there was sometimes ambiguity in answers, or visual issues (for example in geometry) that were much easier to understand via face to face dialogue.

- It is very difficult to read and absorb every single piece of guidance. Examiners reported the training materials were overwhelmingly in written form and not always concise or well edited. For example, materials were reportedly spread across multiple documents, and even in one instance (GCSE sociology) on different websites with different passwords.

- Not all documents are read. This was a strong suspicion among senior examiners, and an admission by some assistant examiners. Failure by examiners to read essential administrative directions was the most common reported evidence of this.

- Senior examiners said they find it 'hard if not impossible' to transfer every piece of clarification and guidance they would give in a face to face session into a readable amount of written guidance. This particularly applied to subjective subjects and subjects with papers including questions requiring extended responses. A small number expressed the view that it was 'unreasonable' to expect examiners to learn entirely online and that not all individuals were capable of it. They said they felt sorry for new assistant examiners trying to learn the entire theory of marking on-screen 'from scratch'.

- There is a lack of opportunity to ask questions during the training stage of standardisation. Assistant examiners said the resulting need to depend on their own interpretation of written guidance meant they had a less confident grasp of the mark scheme and the intention behind the question than they experienced from face to face standardisation.

- Five examiners who had previously standardised face to face independently told us they believed it took them longer to reach the required standard for live marking under online standardisation than it had taken under face to face standardisation.

- The method itself sends signals to assistant examiners about how they are expected to examine. Some new assistant examiners said they had formed the impression from online standardisation that they are intended to 'cope alone' and perhaps not refer issues to their team leader.

- It is harder for the team leader and assistant examiner to build a strong relationship because they do not meet face to face. Some new assistant examiners did not fully understand the team leader role and that they were able to contact their team leader for as much personal advice and clarification as they needed and to escalate 'problem' scripts for review. If the team leader failed to actively encourage dialogue with the assistant examiner (reported in seven or eight cases in this sample) then this dialogue would never take place.

- New examiners found online standardisation difficult because they did not know what to look for in the materials and what pieces of guidance were most essential. Two or three of the nine most senior examiners in the sample (Chairs of Examiners and Chief Examiners) claimed (without prompting by the interviewer) that assistant examiners who have only ever standardised remotely do not get off to a good start, may not wish to stay in the system and may not develop into confident examiners.

> *I notice that the ones who went through the system when it was face to face and paper copies are far more confident with the whole process, it is new* [assistant] *examiners who we have the problems with.'*
>
> Chief Examiner, subjective subject (subject title withheld for reasons of confidentiality)

We conducted follow up interviews with new assistant examiners who had only ever standardised online. At least half felt they were coping well with their papers but when asked for improvements, they all said that they would at some point like face to face training or 'more feedback'.

Nine senior examiners of the 17 in the sample, particularly those leading psychology, sociology and English, expressed concerns about the effect of increased online standardisation on the team leader system. They believed team leaders were being asked to do significantly more under online standardisation than had been expected under face to face standardisation. For example, that team leaders gave detailed advice during the standardisation period, in effect 'individually standardising each assistant examiner by phone'. Some senior examiners also believed team leaders had to mark more standardisation scripts because it took longer to standardise assistant examiners. However, most of the team leaders we interviewed did not say the amount of advice they had to give to assistant examiners was unmanageable, although some did feel it had increased - sometimes significantly.

Around half of the 17 senior examiners that we spoke to said they feared that a small number of very conscientious, usually 'older' and 'experienced', team leaders were 'carrying the system' by consciously compensating for what they experienced as the deficiencies of remote standardisation. These team leaders made proactive telephone rather than just email contact, paid extra attention to new examiners and put time into building relationships to encourage discussion of the paper. These senior examiners believed the changed team leader role had not been formally recognised or addressed by the exam boards in job descriptions for new team leaders and that there was therefore a risk that when 'older' team leaders withdrew or retired, the system would begin to fail.

Examiners noted that under online standardisation they had limited opportunities to comment upon the mark scheme. However, Pearson Edexcel's approach to online and online forum standardisation was praised because it requires assistant examiners to review practice scripts and submit comments on the mark scheme *before* the senior examiners meet for their pre-standardisation sessions. All Pearson Edexcel assistant examiners in the study welcomed the fact that their questions and observations were sought and then incorporated into the guidance and revised mark scheme

Online forum standardisation

Pearson Edexcel is the only exam board that uses online forum standardisation to standardise a significant number of examiners. Under this approach, examiners take part in an online forum that follows the format of a face to face standardisation meeting, but with voice and online chat replacing face to face interaction.

Senior examiners reported that this method is mostly used on 'smaller' papers with few examiners and said it was technically challenging for larger papers. They said that in the past the forum software had not always worked and there were delays during the sessions.

When this method was described to examiners with exam boards other than Pearson Edexcel the majority feedback was that if face to face standardisation was not available then this method sounded preferable to online standardisation. The key reason for this view was the ability to ask questions, albeit in what examiners perceived to be a less convenient way than face to face. However, senior examiners with experience of online forum standardisation did have the impression fewer questions were asked in forum by comparison with face to face. Assistant examiners said working from home, and using a relatively taxing communication method tended to mean that they were keen not to extend the session and so restricted their questions. In face to face sessions they said they were more engaged and that 'you are there anyway so you might as well ask.'

Some senior examiners and team leaders said they were able to identify assistant examiners not contributing to the online forum and either draw them out at the session or follow up performance later via on-screen monitoring, but this practice was not widespread.

On-screen marking

Under on-screen marking students' handwritten scripts are scanned then marked on-screen using electronic icons and basic annotation tools.

On-screen marking for less subjective subjects and papers with lower tariff questions was preferred by most examiners for its convenience and by senior examiners for the greater opportunity it gives exam boards to detect inaccurate marking.

However, although most examiners acknowledged the superior quality control potential of on-screen marking, many senior examiners in charge of subjective subjects and subjects with papers requiring extended responses to questions strongly resisted on-screen marking. These senior examiners did not feel the online software was sufficiently developed to cope with these paper types. There was a consensus among senior examiners that A level papers including single questions worth 30 to 40 marks should continue to be marked on paper until on-screen marking technology improves significantly. Almost all assistant examiners of subjective subjects and subjects with papers requiring extended responses to questions also wanted to continue marking on paper until there were significant technical improvements to on-screen marking software. The main issues identified with the current software were:

- Examiners could not easily scroll backward and forward through a response to maintain an overview of structure and argument. This made it hard to follow complex arguments.

- The viewing panel for text was 'too small'.

- Annotation software was described as 'clunky' and allowing only limited scope for annotation.

Some examiners said they had resorted to printing scripts, but for most this was seen as expensive and unsustainable.

Most teachers spend their year marking hand written student work with pen and paper so examiners said they found adapting to on-screen marking for a brief period once a year difficult. Only two examiners of the 50 interviewed said they had successfully adapted to marking long scripts on-screen.

Another issue raised with on-screen marking was that the exam boards' systems for monitoring on-screen marking (which in many exam boards includes the functionality to temporarily stop examiners from marking) can seem arbitrary or inadequately explained, and that the communications from these systems could be impersonal and insensitively phrased.

Examiners said that the use of red and green lights, vocabulary about 'failure' of seeds and loud alarm noises can add to an assistant examiner's feelings of insecurity and 'panic' if they are already developing feelings of loss of confidence.

Assistant examiners who had lost confidence reported that they had on occasion begun to forget about evaluating student responses against the mark scheme and started to develop a mental attitude of 'battling' the seeding or on-screen marking system.

Examiners are very aware of the existence of tolerances (often defined as the legitimate range within which their marks are allowed to fall) and senior examiners observed that this means some struggling assistant examiners using on-screen marking develop a strategy of marking in the middle of the mark band, awarding neither top nor bottom marks, in order to avoid failing seeds.

Examiner perceptions of the role of the team leader

All assistant examiners agreed that access to a team leader to answer questions and clarify misunderstandings throughout the marking process was crucial if they were to mark accurately. The speed of response from a team leader was seen as particularly important. Assistant examiners, particularly newer ones, said that they valued reassurance from their team leaders. They particularly appreciated team leaders who communicated proactively with them and who explained aspects of the process that were unfamiliar or potential sources of stress – for example, outlining what he/she was looking for when sampling the assistant examiner's marking, and explaining the reasons why an assistant examiner was stopped from marking specific questions.

Examiners reported great differences in team leaders' communication with and support for assistant examiners. The majority of team leaders received praise from their assistant examiners, as 'brilliant' and 'supportive' although a proportion were not felt to be either adequately available or supportive. While most assistant examiners had several phone calls or long email exchanges with their team leader there was more than one instance where examiners said they had only received one introductory email from their team leader during the entire examining process. Team leaders said this could happen when an experienced assistant examiner has been marking well on a low tariff paper and does not need intervention. However as the assistant examiner was not usually made aware of this fact, the lack of contact was commonly interpreted as lack of interest or forgetfulness rather than a sign of their own marking excellence.

The contact made with assistant examiners by team leaders varied widely. Examiners reported that, across all exam boards, team leaders are instructed to contact each assistant examiner at least once at the beginning of the marking process, but the method of contact is not specified.

Examiners said that, following the initial contact by the team leader, there were no best practice guidelines on the frequency, tone and method of team leaders' contact with assistant examiners. One senior examiner mandated team leaders to contact by telephone in order to build relationships but other team leaders used solely email.

Examiners also said the exam boards gave no recommended response times for team leaders to respond to assistant examiners' queries, although most team leaders said they tried to respond to assistant examiner queries quickly because as examiners themselves, they understood the 'need for speed'.

Assistant examiners reported team leaders can either be patient in their response to queries, or brusque and appear to have little time. Assistant examiners said that an impatient attitude could discourage them from asking questions, as could their own awareness of their team leader's high workload.

> *'Both my team leaders to date were teaching at school and were also team leaders they didn't have much time. They did 'make time' but I always felt I didn't have enough time to really explore stuff. I could have, but I felt I shouldn't. I felt I would have been intruding on their time.'*
>
> Assistant examiner, English literature

Assistant examiners who had struggled to mark within tolerance said they believed they had not made adequate use of the team leader. They said that in retrospect they believed the best way to operate was to make very intensive use of their team leader.

> *'It is better to pester* [the team leader]*. I used to get embarrassed by how annoying I was being and sometimes I thought it wasn't welcome, but it is better to understand every single little tiny bit of the mark scheme and ask all the stupid questions. It is better to do that than **you** get blamed for getting it wrong.'*
>
> Assistant examiner, sociology

Some new assistant examiners were confused about the role of the team leader and how to use them. Some were restricting their use and some were not sure what kind of queries they should pass up to their team leader or what level of support to expect. Those assistant examiners standardised online did not feel the induction packs made this very clear and three complained the team leader contact details were not easy enough to find.

Examiner perceptions of papers and mark schemes

Examiners had fewer comments on mark schemes than teachers. However newer examiners in particular reflected teachers' comments that there was too much internal jargon and "vagueness" in mark schemes. In particular they made similar criticisms of language relating to level descriptors (for example, defining "sophisticated" or "good").

We encountered some examiners who had marked for several exam boards. These examiners were able to draw some comparisons across exam boards but felt that the mark scheme was more a reflection of the specific philosophy of the paper than of the exam board.

Senior examiners said there are always two versions of the mark scheme:

1. Pre-standardisation mark scheme: written alongside the paper and issued before standardisation begins. Some senior examiners, aware that this mark scheme will later be modified, said they often write this in a 'fairly generic' style similar to that of the mark scheme published with the specification.

2. Post-standardisation mark scheme: issued when the standardisation process has been completed. This is, in the senior examiners' view, the 'real' mark scheme because it is fully related to what students have written in the paper and has been commented on by a range of other examiners.

Six of the senior examiners expressed concern (unprompted by the interviewer or discussion facilitator) that assistant examiners may print off the first mark scheme, start to work with it, then pay insufficient attention to the final mark scheme and guidance or mark most scripts using the first scheme.

We are asked to note that Ofqual's Code of Practice[16] specifies that exam boards must issue the post-standardisation mark scheme to all examiners to ensure that all examiners have the definitive final version of the mark scheme. Assistant examiners must not finalise any marking until they have received this authorised revised version

With face to face standardisation the style of the written mark scheme was regarded by examiners as less important as there is discussion around it at the standardisation meeting. However, when an assistant examiner is working through a mark scheme alone via remote standardisation the clarity of the mark scheme and the language used within it becomes very important.

---

[16] Ofqual (2011) "GCSE, GCE, Principal Learning and Project Code of Practice".
http://ofqual.gov.uk/files/2011-05-27-code-of-practice-2011.pdf

Examiner perceptions about marking subjective subjects and subjects with papers requiring extended responses to questions

Examiners agreed with some of the issues suspected by teachers around examiners' lack of familiarity with some texts and topics. They said there was a particular issue if assistant examiners attempted to mark English literature texts they were unfamiliar with. It was reported that examiners on some papers have to sign a paper verifying they are familiar with the set texts before they are accepted to mark. Senior examiners told us that if an assistant examiner did feel unfamiliar with a text in their view it was acceptable or even expected that they should pass it to the team leader. However, in practice the (seven) English literature examiners we spoke to had not in the past believed this was a permissible use of the team leader.

A senior examiner in English literature said:

> *'We used to give examiners notice of what books 'their' schools had studied so that they could prepare for their examining. Now we have a situation where people can be asked to mark on anything* [random allocation via on-screen marking] *- even if they haven't read it for years. There is a reject option if an examiner feels they can't mark a script but I feel some are 'having a go' when their own understanding of a text is not much better than the candidate they are marking.'*

One new assistant examiner accepted an additional allocation of GCSE scripts which she found included unfamiliar texts. As it was late in the process she was asked to "get familiar" with the texts. She felt that although she had worked hard to do so in her view her final marking was not of the standard she would have liked.

Three examiners of psychology and sociology said they had to deal with scripts that included topics unfamiliar to them within their total allocation. In psychology or sociology the student usually has to quote cases relevant to the topic. Some examiners said if a case was unfamiliar to them they would try to look it up online or discuss with their team leader. However senior examiners said they occasionally see instances where cases have been credited incorrectly, or not at all by assistant examiners. The risk of this was felt by assistant examiners to be particularly acute if the examiner was behind on an allocation and running out of time.

Other problems quoted by examiners were:

- Deciphering poor handwriting, particularly on long essays (this was described as a very common problem).

- Lack of guidance about what to do about unusual or ambiguous responses that might or might not fit the mark scheme (for example when the mark

scheme said 'or any other relevant answer'). This was a problem that particularly affected the newer examiners in our sample.

- Instances where the mark scheme and example scripts provided did not give adequate guidance for mid-range students or for students falling in the middle of a mark band (rather than the top or bottom of the mark band). This was less often spontaneously reported but was a consistent issue on prompting for over a third of examiners who mark papers including questions requiring an extended response.

The issues were particularly likely to apply to subjective subjects and papers including questions requiring an extended response, but also applied to other subjects and paper types.

Difficulties and challenges that examiners experience in marking

Most examiners in our sample were prompted to examine by a desire for professional development and a wish to improve their students' results through a better understanding of the specification.

Of the remainder those who were retired teachers were motivated largely by a desire for mental stimulation, social contact with like-minded others and retirement income. Only a few examiners, exclusively younger teachers, university lecturers, and education consultants in this sample, started examining mainly because they wanted to earn additional income.

All examiners took their responsibility 'to get it right for young people' very seriously, believing their task to be very important. Senior Examiners expressed an opinion that most examiners are conscientious by nature.

Only five of our sample of 25 assistant examiners reported enjoying examining without prompting from an interviewer. For the majority in our interviews and group discussions, processes such as online standardisation and on-screen marking mean that examining is becoming less rewarding but they feel they are still learning something that makes the process worthwhile. However, there remained four or five examiners who found the process stressful, whether unrewarding, 'isolating' or in a few cases actively damaging to their self-esteem (for example, if they performed well for a period of time but then began failing seed items or receiving a lower overall performance score at the end of the marking process).

The chief challenge of marking for most examiners is that their allocation of papers has to be completed 'against the clock'. A number of scripts or items are allocated to them and have to be fitted into a set 'marking window'. Many examiners said they feel a continuous level of background anxiety while they are marking.

Assistant examiners can unexpectedly fall behind and this causes anxiety. This can happen, for example, because:

- Examiners find themselves unexpectedly 'too tired' to mark with the necessary accuracy of judgment for the hours they have set themselves that day. This is a particular issue for teachers and parents of small children.

- Some scripts take longer to mark than others.

- There are delays: at the beginning of the marking process, standardisation results and clearance to mark may come through 'too slowly', particularly with pen and paper marking. During live marking an assistant examiner may be 'stopped' for marking inaccuracy and not re-started within what he/she sees as a reasonable time, or a team leader may not get back to the assistant examiner promptly on a query. If an assistant examiner is marking on-screen, an electronic marking system might 'go down' or be cumbersome in its design, making each paper take longer to annotate than the assistant examiner predicted.

- There can be administrative issues such as mistakes or ambiguity in the instructions issued by exam boards. These issues, reported by four or five assistant examiners, are particularly resented and assistant examiners feel that their pay is not sufficient to compensate them for these kinds of inconveniences.

We observed differences between assistant examiners marking papers with lower tariff questions and those marking papers requiring extended responses to questions.

Those assistant examiners marking papers with lower tariff questions (for example, GCSE maths and science papers where each question is worth no more than four marks) had a more pragmatic attitude to marking exam scripts. The task of marking these papers is a relatively controllable one for them and assistant examiners can estimate the time it will take quite accurately. There are fewer and less difficult independent decisions to make, rendering the task less stressful and onerous. The 'value equation' in terms of the work put in balanced against the rewards of examining is therefore seen as fair by these assistant examiners and pay can function as a stronger motivation:

> *'It is only a few weeks at the end of the summer when many of my students are on study leave, and it pays for my holiday.'*
>
> Assistant examiner, business studies

In more subjective subjects and subjects with papers requiring extended responses to questions (where 30 or 40 marks are available for one question) assistant examiners felt more anxiety and thought they needed to put in more effort than assistant examiners marking more straightforward scripts. Assistant examiners felt they carried more individual responsibility to 'get it right for the student' in those papers that required the examiner to make subjective judgments.

The time required to mark each response to a question is also less predictable in papers with questions requiring extended responses. An exceptionally long response, poor handwriting or an unusual response which the assistant examiner is not sure is admissible can extend the time needed to mark a single paper. Judging the correct mark band and particularly the correct mark within that band can take time, particularly with mid-range ability students.

Some assistant examiners expressed more worry about their judgments than others and said this can affect the length of time marking takes, and the stress and tiredness they feel. Most assistant examiners were aware that they could not mark accurately when tired and therefore should not continue marking. However those assistant examiners then said they felt the pressure of time even more acutely when they resumed marking.

Across all paper types assistant examiners often expressed feelings of insecurity, and some said that they take indications of 'failure' very personally. Many team leaders reported that assistant examiners need frequent reassurance that they are performing correctly, particularly during the standardisation or early qualification period.

Senior examiners viewed providing reassurance as a priority because they believed examiner confidence is important to accurate marking. Senior examiners held a theory (based, they said, on long observation), that a 'shaky grasp of "the standard"' and consequent lack of confidence can lead to a downward spiral of performance in assistant examiners. An already worried assistant examiner can become less and less secure about their grasp of what is required, how to award marks and even what part of the guidance they should treat as a priority. Senior examiners believed this can lead to erratic marking. Senior examiners also said that this lack of confidence may cause assistant examiners who are marking on-screen to mark over-cautiously - that is, marking to the middle of the mark band so as not to 'fail' seeds.

Assistant examiners reported that they believe their marking is faster and more accurate when they are confident in their grasp of 'the standard' and the 'intention' of the paper and the questions.

How marking of exams might be improved

1. Return to face to face standardisation for some paper types

As outlined at the beginning of this section, the majority of examiners of subjective subjects and papers with questions requiring extended responses felt a return to face to face standardisation for these subjects and papers would improve both marking and rates of examiner retention.

2. Introduce mandatory face to face training for new assistant examiners

If face to face standardisation for subjective subjects and papers with questions requiring extended responses is not possible, almost all examiners felt initial face to face training for new assistant examiners and assistant examiners on new specifications would improve their long term quality of marking. Senior examiners felt this training would get new assistant examiners off to a better start and make it easier to standardise these examiners online in future years.

The type of training assistant examiners wanted was very similar to the content of a face to face standardisation session: practice, alongside other examiners, in marking scripts and applying the mark scheme. Assistant examiners believed that in order for training to be effective it had to be delivered by someone very confident about the specification and the mark scheme, preferably the Principal Examiner. They also felt that it would be useful to go through recent papers.

Assistant examiners also said they needed much more notice of training than they were currently given. The current exam board training for new assistant examiners, if offered at all, tended to take place at 'very short notice' and 'far too close to the exam.' This had prevented attendance for some examiners who wanted to attend.

3. Review and redefine the role of the team leader

Senior examiners wanted exam boards to address the risk they perceived of the system becoming over-dependent on team leaders contributing above and beyond the call of duty.

This included reviewing how online standardisation impacted the team leader role, accepting the need for a revised team leader job specification and pay if necessary and producing best practice guidelines for managing examiners, preferably based on the experience of the most experienced team leaders. They suggested the role would then be delivered with more consistency.

Assistant examiners suggested that exam boards should provide guidance around the frequency and nature of team leaders' communications with the assistant

examiners in their team. They also wanted the opportunity to give feedback on their team leaders that would feed into the team leader's performance reviews.

4. Improve the retention and development of examiners

Many senior examiners firmly believed the examiner base was beginning to polarise between older (retired) examiners and young Newly Qualified Teachers (NQTs) or Recently Qualified Teachers (RQTs) who examined for a year to get a 'quick fix' on the specification.[17]

Some Chief Examiners and Chairs of Examiners had the impression they were losing experienced examiners to retirement or the 'smaller' exam boards because of growing use of online standardisation and examiners' consequent declining job satisfaction. They also thought they were failing to retain young examiners after their first year or two. This was upsetting for them because in their view these examiners had 'a lot of effort put into them'.

Additionally, senior examiners felt that there were too few experienced older teachers and heads of department examining regularly (although Ofqual's quantitative research suggests this is not the case).

Senior examiners would like to see these issues addressed through the exam boards' recruitment strategies and also through greater Continuing Professional Development (CPD) activity to retain and develop examiners, particularly new ones.

5. Give examiners more feedback

Most examiners wanted more feedback and reassurance on their performance than they currently receive from their exam boards. The score they receive from the exam board is focussed on the exam board's internal statistics for accuracy and timeliness. Some examiners do not receive their grade directly. Those who get their grade say they do not find it transparent and they cannot learn from it:

> *'Personally, I believe that I would benefit from direct and statistical feedback following each marking session, knowing how my marks*

---

[17] The findings from Ofqual's online examiner survey indicate that most examiners are experienced teachers: 62% of the surveyed examiners said they are current teachers or lecturers, 99% had teaching experience, and fewer than 1% said they had less than three years' teaching experience. However, it was not possible for the survey to measure the extent of any changes in the demographic profile of examiners in recent years. For further details see "Review of Quality of Marking in Exams in A levels, GCSEs and Other Academic Qualifications. Findings from survey of examiners, May 2013." (Ofqual, 2013)

> *differed from the norm, how many of my scripts are returned for review, etc.'*
>
> Assistant examiner, geography

Examiners said they want to feel they are doing well and to receive feedback and encouragement. New assistant examiners frequently claimed they were ending their first season of examining feeling that they had not done particularly well. They might have been stopped a few times in on-screen marking, and their team leader might not have contacted them during the process to encourage them or at the end to reinforce their confidence.

Some senior examiners also felt there was a 'wasteful' attitude in exam boards which meant examiners who made mistakes for example in consistent over or under marking were too rarely offered remedial training. They also worried that the way in which on-screen marking systems 'locked out' examiners was sometimes peremptory and that the system sometimes lacked an element of human judgment. Senior examiners believed new assistant examiners need some input to develop well and confidently and that even experienced examiners have 'wobbles'. They believed that assistant examiners' temporary issues with marking accuracy should sometimes be addressed by retraining, particularly if a new assistant examiner shows initial promise or if a more experienced assistant examiner has previously performed well.

6. **Give assistant examiners more help in marking 'obscure' texts and topics.**

Examiners put forward several solutions to assist more accurate marking of these:

- Give the marking of rarely-sat texts and topics to examiners who know them well, and encourage examiners to pass up texts/topics they are not familiar with.

- Where that is not possible, give examiners more lead time when inviting them to mark, so they have sufficient notice to read set texts.

- In subjects such as psychology and sociology, set fewer topics overall so examiners do not need to be familiar with so many.

- Provide online reference systems for sociology and psychology research studies and support materials for rarely-sat topics so examiners can check answers quickly.

7. **Introduce visual methods and new technology to complement written guidance for online standardisation.**

Examiners suggested the introduction of team leader video conferencing with functionality that allowed examiners to view the diagram, visual example (geography), script or text the speaker was referring to. This was felt to be particularly helpful for maths diagrams.

More and better use of video to complement online standardisation: for example videos of face to face standardisation sessions, outlines of key procedures, or senior examiners explaining the intention of the paper and questions.

6. Attracting more examiners into the profession

Teacher views of examining

On the whole examining has a reputation among teachers in our sample as an unenjoyable slog that affects personal life negatively. Teachers who don't examine can be afraid to take it on. Their apprehension is caused by anticipation of a high number of scripts to be marked in a very brief window at the beginning or end of a full working day or at weekends. The experience of marking within this 'window' is imagined to be stressful and unpleasant. Some established teachers felt that unless you were 'super-organised' examining would be almost an impossible task.

A great deal of this impression comes from peer group feedback. Some capable and subject focussed teachers who are long term examiners (often becoming team leaders) claim to enjoy examining, and some senior teachers strongly advocate it to their departments. However, a large number of serving teachers who examine report that their experience of the brief marking window period is not pleasant, the effect on family life is 'isolating' and that they only examine for the wider professional development or income benefits.

This poor impression is further reinforced for teachers by a perception of a shortage of examiners, which can give an impression that 'nobody wants to examine.' Examiners felt that exam boards' approach to recruitment advertising amplifies this impression and fails to raise the perceived status of examining or examiners. Recruitment advertising was sometimes perceived to be last minute, suggesting to teachers that there was a significant shortage of examiners in those subjects.

It also appeared to teachers that recruitment is not targeted specifically at senior classroom teachers. Many had the impression (perhaps incorrectly) that a requirement for specific qualifications was not mentioned in advertising.

Teacher perception of examiners is polarised, which means that the role had an ambiguous or contradictory status and image when probed in discussion groups:

- On one hand, senior teachers or department heads who examine get peer respect, their opinions are sought and deferred to and they can obtain 'guru' status in school, often being consulted on exam technique beyond their subject. A junior teacher can reportedly get a career and peer group 'leg-up' from examining.

- On the other hand, examining can be positioned by school leaders as 'free INSET' (In Service Training), 'good for NQTs for a year or so', useful for teachers earning lower salaries, a good way for a teacher in a non priority subject to get some CPD, or a cheap way for the school to try out a new specification.

Teachers' views of examining

The attitude of the school and the type and number of teachers who consequently examine in that school has a large effect on teachers' desire to become an examiner. Teachers do regularly take an independent decision that they want or need to examine, but the school's attitude can also prevent teachers doing so.

Some schools are not only positive towards, but highly strategic in their attitude to examining. We observed a highly strategic use of examining in particular in the selective state and the selective independent schools in our sample. In these schools the head teacher reportedly saw examining as a clear and direct route to the improvement of academic results: 'examining is the best INSET'. These schools saw examining as delivering continuous improvement of results:

> *'I would be frightened to stop examining now. Being part of the process has given me such a deep insight into the paper I teach…we feel with all the changes to the exams, it would now be so dangerous for your students to miss out on this insider information.'*
>
> Teacher, independent school

The independent schools in our sample were most often testing alternative qualifications to A levels and most of the teachers involved said they would never take this risk without first taking the opportunity to examine and get 'inside the qualification.'

Strategies employed by independent and selective state schools to use examining to understand qualifications included:

- Encouraging at least one member of every department to examine.

- In one selective independent school (London) a number of department members were between them examining every unit on an A level history paper.

- Preferring examiners over non-examiners when recruiting new teachers.

Exposure to the thinking methods of the Principal Examiner via team leadership or being part of face to face standardisation was felt to help teachers to guide their students to answer questions more accurately:

> *'By now we* [team] *understand his little nuances.'*
>
> Teacher, independent school

Some teachers also believed being an examiner could help with the enquiry about results process: not because their school is viewed more favourably, but because teachers feel they are more likely to be able to spot rogue or outlier results and go back with the right type of enquiries within the most effective timeframe.

Other schools are not as convinced or aware of the potential benefits of examining.

Overall, the potential of examining to improve both results and teaching is felt by examiners to be under-estimated in some schools, particularly maintained schools:

> *'I was in a state comprehensive* [before my current job] *I think if* [my previous school's senior leadership] *understood more about what benefit it can give to grades there would be more emphasis and encouragement. I don't think in my experience people were that clued up to it. People did it round school but a handful and no one knew if you did or didn't and there was never any suggestion that doing so was helpful to improve grades.'*
>
> Teacher, selective state school

Examiners in schools with a more negative or passive attitude to examining do not feel their status is an asset and may actively keep a low profile, concerned that it is a disadvantage to them.

Some examiners in the London group discussion who came from schools that were quite negative about examining were surprised when examiners from more positively disposed schools encouraged them to put examining 'on their CVs'.

We also encountered some senior and experienced teachers, including in two cases in the independent sector, who felt examining was slightly beneath or 'behind' them. That is, that being asked to 'mark', or having to do it, was an inconvenience they had escaped by virtue of seniority or their higher levels of pay. They might see it as more appropriate to encourage a younger department member to take it on.

Reasons why a head teacher may refuse a teacher permission to apply to become an examiner include:

- A view that marking is onerous and desire not to add additional stress to his/her staff.

- Concern that examining would take the teacher away from other roles and responsibilities at the school.

> *'It is all to do with resources and how you deploy them. We have greater government control. We have to balance all the juggling balls. In the private sector it is all about results. In the state we have a lot of things to*

> *report on. Pupil premium, my maths department yes I may want an examiner in there but I also want* [a teacher to carry out] *an intervention,* [a teacher] *who is going to deal with pupil premium students, I might want a* [teacher to specialise in] *gifted and talented* [students] *because all these things are things that my Governors and Ofsted will dig into, we balance many things. That is a fact of education today.'*
>
> Head teacher, comprehensive school

- Concern that students' learning would be affected if the teacher needed to be released for examining duties during term time.

- Some head teachers said they had experienced teachers becoming more interested in their examining role than their teaching role, getting 'sucked in' to the examining system, requesting frequent absences and then failing to share their learning across the school. These examiners were usually senior examiners, who take more time out to attend pre standardisation. However, this experience affected these head teachers' attitude to all subsequent requests to examine and head teachers often over-estimated the number of days an ordinary assistant examiner would need to be released from teaching.

Attracting more teachers into the profession

There is a latent belief that examining does improve results which teachers felt could be tapped into and exploited by exam boards looking to recruit new examiners. All teachers interviewed agreed examining was by far the most direct route (sometimes the 'only' route) to understanding how marks are awarded in external exams and improving performance.

Teachers who also examined reported they found out things they could not gauge from the published specification or even that contradicted the specification:

> *'What I learned by examining is actually retrograde to the specification. The English literature unit I examine is (nominally) highly weighted on context so we teach that and a lot of time is spent on that, encouraged by the spec. But the examiner* [at standardisation] *said 'I know it is very heavily weighted, but actually, don't worry about it!'…That is an advantage to me, I have gone back to the department and said 'don't worry about that as much now, but there are other schools who do not know that.'*[18]
>
> Teacher and examiner, comprehensive school

---

[18] This view was repeated by two other English literature examiners in our sample.

Two examiners teaching GCSE media studies and sociology said they had learned and passed on exam techniques to exploit positive marking which had improved student confidence and boosted the school's grades.

For these reasons, the majority of teachers agreed with the suggestion that exam boards should emphasise the possible benefits of examining to a school's exam results to prospective examiners.

Teachers also recommended that exam boards should:

- **Do more direct outreach with schools**. Various forums were recommended. For head teachers, routes suggested were The National College for School Leadership, Head teacher unions and local and regional clusters, academy management groups, regional subject groups and local authorities for maintained schools.

  The IB's outreach programme was praised by all the teachers who taught the IB for instilling a 'we are all in it together' spirit in its schools which had encouraged schools to release examiners who were teaching the specification. There was also strong promotion of examining in training sessions run by the IB for teachers.

- **Consider changing the way they make payments to schools** where examiners need to attend meetings during term. One head teacher in a comprehensive school said their school would prefer the money paid to the school to reimburse the cost of a supply teacher to be paid differently for example, a cash sum for the total number of examiners the school was providing rather than offering individual cover vouchers. The head said a predictable lump sum contribution would have been more flexible for them since they did not find cover vouchers particularly useful, particularly at a time when they needed more senior teachers for exam preparation.

- **Make it easier for examiners to share their gained expertise in school** with 'free' or additional sample scripts and practice material.

The relevance and role of pay in encouraging teachers to examine

The perceived fairness of examiner pay depended on the difficulty of the task, the length of time taken to mark a paper, any collateral stress (for example from mistakes in exam board administration or technological breakdown) and finally, on the relative income of the examiner.

The pay from the main UK exam boards was not seen as particularly good *per se* but it was more often seen as good value for papers with short questions with few marks and unambiguous answers. However, pay by the UK exam boards seemed, on the

basis of what examiners told us, to be sometimes similar for a GCSE and an A level script.

As examiners told us that they felt the effort involved in marking papers with questions requiring extended responses is greater than the effort involved in marking papers with predominantly lower tariff items, this can lead to GCSE examiners (who in most subjects are more likely to be marking papers with predominantly lower tariff items) feeling they are very well rewarded, and A level examiners (who tend to mark more extended responses) feeling relatively underpaid.

Senior teachers on the higher tax rate feel relatively less incentivised by the pay.

> *'£3 a script is ludicrous for a child's future… that is what £5 gross comes out at if you are a higher rate tax payer. You pay more,* [examiners] *feel more responsibility … I suggest £20 an hour equivalent would be appropriate.'*
>
> Head, independent school

Some senior teachers in the study suggested that if the exam boards wanted more senior teachers to examine they should follow the teacher pay scale and pay a higher rate to more experienced teachers.

Overall, examiner pay, which is not seen as high, looks less attractive if the examiner is not getting the other benefits they want out of marking. These include social, professional, self-esteem and school improvement benefits. Pay is a small part of the value equation for a teacher who wants to end the exam season feeling good about themselves and having had a rewarding or professionally useful experience. Where stresses multiply and personal rewards decline then the value equation of examining can simply cease to add up for some examiners.

Appendices

Appendix A - Methodology

Our research structure included group discussions and depth interviews:

- Group discussions allow for debate and discussion, which is important as it reflects how matters are discussed between peers in reality. Debate can jog memory: prompting additional examples from memory and helping respondents be more creative and imaginative about answers to questions such as 'what could make you more interested in examining?'

- Face to face depth interviews were used for head teachers to reflect their status and provide a more confidential context for conversation.

- Telephone depth interviews were used for teachers of harder to find qualifications and for examiner follow ups so that we could target individuals who were relevant wherever in the country they taught or worked.

The following table shows the composition of our sample, and the number of participants who teach/examine subjective subjects and subjects with papers requiring extended responses to questions.

Sample NB: some senior examiners held multiple roles and titles	Number Included in the sample	Of which subjective subjects and subjects requiring extended responses[19]
Head teachers who were not examiners	4	n/a
Teachers who were not examiners	28	18
Chairs of Examiners	5	4
Chief Examiners	4	4
Principal Examiners	8	7
Team Leaders	8	6
Examiners	25	16
Total	**82**	**55**

We identified the more subjective subjects by looking at the subjects with higher levels of Enquiries About Results, and took guidance from Ofqual about the subjects

---

[19] The subjects over represented in the sample were history, psychology, sociology, politics, business studies, classical civilisation, International Baccalaureate (IB) Theory of Knowledge, IB Extended Essay, geography, religious studies and English. The paper types over represented were A levels with longer 30 and 40 mark essays and also a proportion of GCSEs with essay elements, for example English literature, sociology, business studies, media studies and religious studies.

accepted as requiring extended responses to questions that should also be included. Our sample also included teachers and examiners of other subjects (such as maths and science) to represent the views of teachers and examiners of the remaining qualifications.

Teacher sample: considerations

The teacher sample included teachers from the following backgrounds:

- A range of teacher experience and role from NQTs through to experienced teachers, heads of department and head teachers.

- A range of school types: independent, selective state, state comprehensive (local authority maintained and academies)[20].

- A range of different sizes of school.

- A range of school success – high and low achieving academically and a full range of Ofsted categories. This is because these schools differ greatly in approach and attitude and we felt there might be differences in the encouragement they gave teachers to examine.

Examiner sample: considerations

The examiner sample was constructed to include the following:

- A range of experience (from examiners with only one year's experience up to examiners with over 40 years' experience) in order to understand the influence of recent changes in standardisation, marking and training on new examiners.

- The full range of roles within the exam boards.

It became clear as the study progressed that many senior or retired examiners hold multiple positions with different exam boards and on different papers. They may even also work for some exam boards as assistant examiners.

Clerical and graduate markers were not deliberately sampled because there is almost no scope for ambiguity in the types of questions marked (e.g. multiple choice or one mark questions). Moderators were also not sought, because this study focused only upon the external marking of exams.

[20] The independent and state selective sample was a higher percentage of this sample than in the population in order to cover off teachers and examiners of newer and lesser-sat qualifications, which appear disproportionately concentrated in these school types.

However because of the very wide experience of examiners in the sample we inevitably spoke to examiners who had experience of all the 'disciplines of marking' for example, revising and moderating and also some who had been clerical markers in the past.

Recruitment

We recruited our examiners from respondents to Ofqual's online examiner survey[21] in April 2013, in which respondents were asked if they would be prepared to participate in further research. The data from Ofqual's survey enabled us to:

- Recruit participants who, on the basis of our analysis of their answers to Ofqual's survey, seemed to reflect the majority views of examiners of their subject in their exam board.

- Identify senior examiners and examiners of the smaller qualifications.

- Frame our questions 'intelligently' to reflect and investigate the common issues examiners raised in the quantitative study.

Teachers were recruited directly from schools to meet the quotas of school type and size, and subjects and qualifications taught. We recruited teachers from within the general teaching community. This was done in the case of our group discussions via a professional fieldwork company and in the case of 'hard to find' subjects, qualifications and exam boards, by researching which schools taught the relevant qualifications and then calling and emailing directly to invite teachers and heads of department to participate.

---

[21] Ofqual (2013) "Review of Quality of Marking in Exams in A levels, GCSEs and Other Academic Qualifications. Findings from survey of examiners, May 2013."

First phase sample: Teachers

Group number	Location	Teacher type	Exam board and exam	Subject
2 x teacher group discussions of 1hr 45 each				
3	London	Teachers in comprehensive schools	GCSE and A level Min of 2 per group experienced in other exam types (IB/IGCSE) Spread of exam boards Min 3 academies per group	One history and one English teacher per group Spread of subjects with one group emphasising subjective subjects/ subjects requiring extended responses and the other reflecting a broader range of subjects
4	Manchester			
4 x face to face depth interviews of 1hr each				
1	South	Head teachers	State selective	
2			Independent school	
3	North		Local authority maintained comprehensive	
4			Academy (comprehensive)	
10 x telephone depth interviews of 1hr each				
1	National spread	State selective	GCSE and A level.	Classics
2		State selective	IB	Theory of Knowledge
3 a /b		Comprehensive and Independent	IGCSE	English
4		Independent	IGCSE	English literature
5		Independent	A level and GCSE	History
6		Independent	Pre-U	English literature
7		Independent	IB and GCSE	Extended essay Modern foreign languages
8		Comprehensive	Newly qualified teacher	Religious studies
9		Comprehensive	Newly qualified teacher	English literature

First phase sample: Examiners

No	Location	Examiner type	Board & exam	Subject
Groups				
1	London	Assistant examiners	AQA, Pearson Edexcel, WJEC, OCR,IBO	Quotas on subjects similar to the teacher sample above
2	Manchester			
Examiner telephone depth interviews				
1	National spread	Assistant examiners	IB	Theory of Knowledge
2			Pre U	English literature
3			IGCSE	History
4a		Principal Examiners	AQA, CIE, IB, OCR, Pearson Edexcel, WJEC	English literature AS/A2
4b				Chemistry
				Psychology
4c				Classical civilisation
5				Religious studies
6				A Level history
7				Geography
8				English literature
9				
10		Team leaders	AQA, Pearson Edexcel, OCR	Sociology
11				Religious studies
12				History

Phase 2 Follow up sample (looking at remote standardisation and team leaders)

Number of interviews	Position	Exam Boards	Subjects
4	Chairs of Examiners	Include AQA, IBO, Pearson Edexcel, OCR	Subjective subjects/ subjects requiring extended responses
5	Chief Examiners		4 x subjective subjects/ subjects requiring extended responses 1 x modern foreign languages
3	Team leaders who are also teachers and standardise online		English, geography, maths
5	Recently appointed examiners (3 years' experience or less) who standardise online		IB Theory of Knowledge, business studies, English, modern foreign languages

Appendix B - Discussion guides
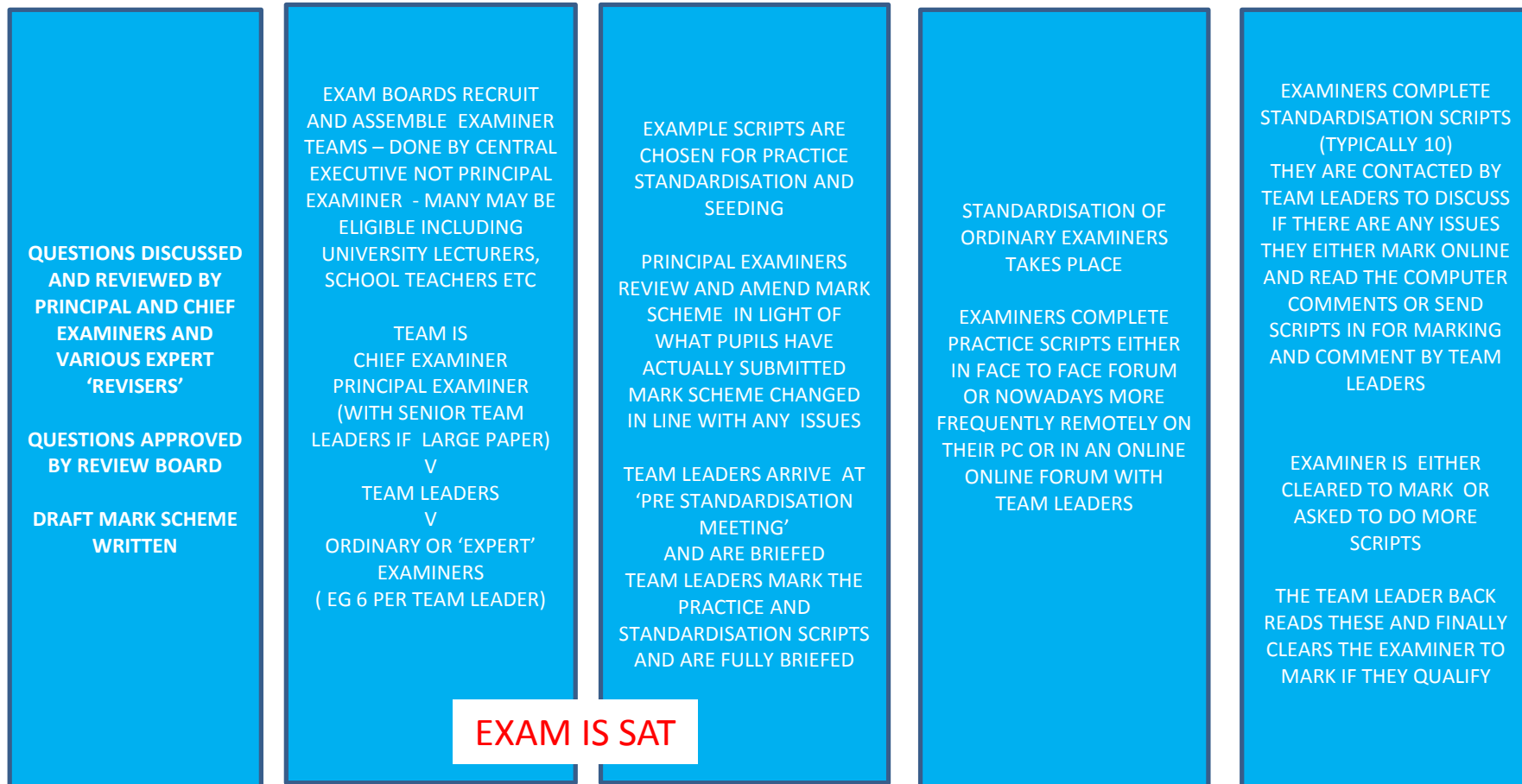
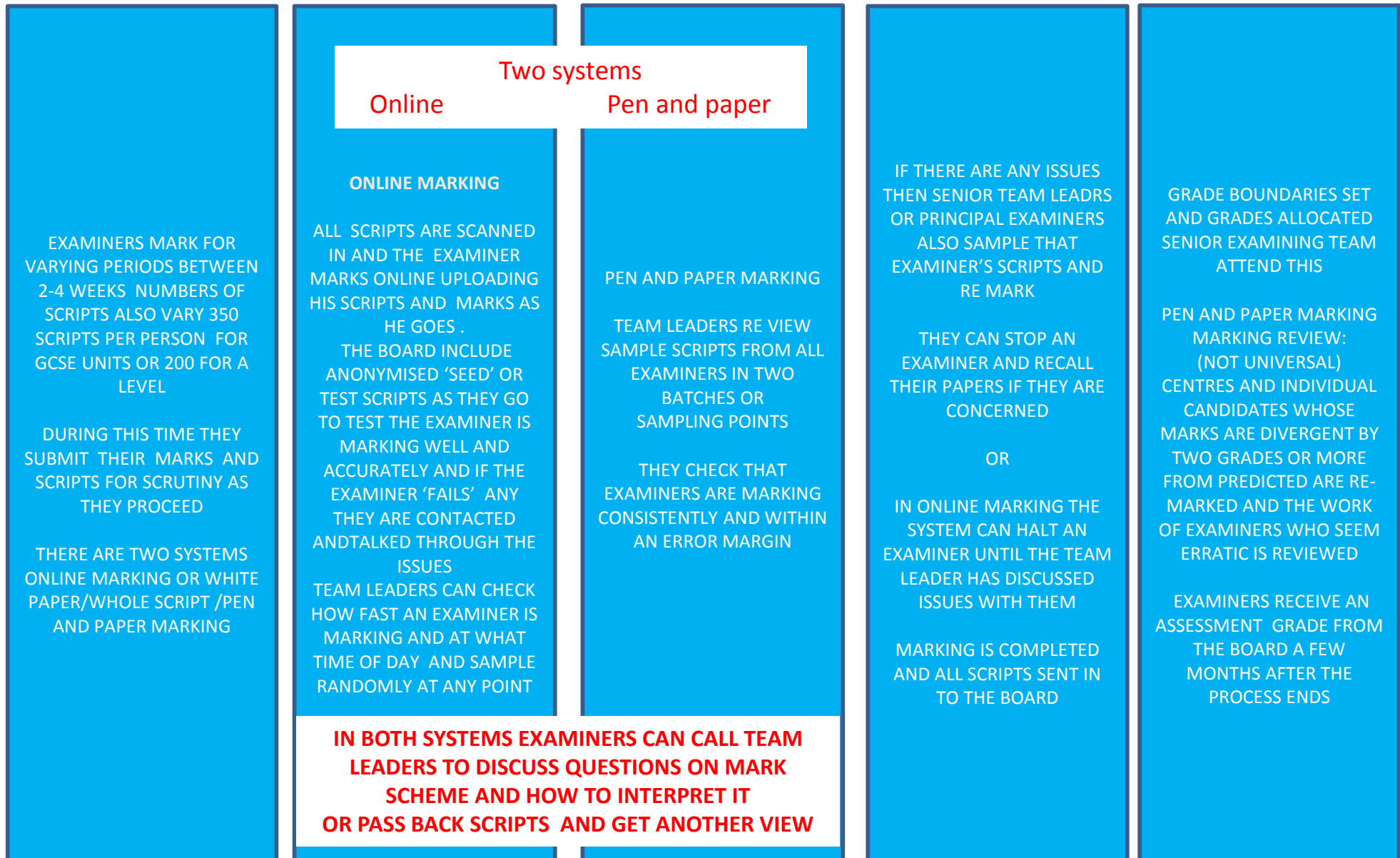|		
	Research Objectives 1. Teachers' perceptions of quality of marking of GCSE and A-level and equivalent exams 2. Levels of understanding of the marking process 3. How marking of GCSE and A-level and equivalent exams might be improved 4. What information teachers would like to receive about marking from Exam Boards 5. How teachers feel about the role of examiner 6. How to attract more teachers into examining work	
<td colspan="2">**Introduction - perceptions of quality of marking**</td>	

	• Explain purpose of the research : this project from Ofqual is focussed on marking not grading and at complex essays not multiple choice – explain that we may occasionally need to cut conversations short if we get off topic but it is not personal as we have a bit to go through.
	• We have to concentrate on the topic in hand rather than other aspects of marking you may feel justifiably passionate about!
	• Even where you feel you can't comment from experience we are still interested in perceptions and impressions and where they come from
	• Self-introduction:
	   o name,
	   o subject teach
	   o board teach
	• What do you define as 'good quality marking' – flipchart
	• Examine concepts of reliability/reliability/consistency: see how big a part of the teacher perception this is
	• Do you feel or observe that marking is reliable or accurate
	   o Why do you feel that?
	• If you feel ever inaccurate then how often? – To what extent?
	• What is an example of reliable /unreliable marking (flipchart): Look at examples and instances they have prepared and brought along of unpredictable marking – probe re-marking etc.
	• Any ideas as to why this might have happened?- hypotheses
	• Good instances? – Why this is? – Any ideas as to why it works well in this example, e.g.; subject or exam board?
	• Impressions/images and pros and cons of Exam Boards particularly in this context
	• Or of exam types in terms of predictability?
	• Or subjects?
	• What bothers them as teachers most and where do they want to see most attention focussed – on the outcomes in terms of consistent marking or in the guidance they are given to prepare pupils?
	• *What drives their perception of marking- (moderator probe*
		40 Mins

	this continuously as we go through – is it other teachers' experience, grading, enquiries about results, the media, knowing examiners)	
How much do teachers know about external marking?		

	• What happens to papers once they leave the school (write this down – any ideas, anything they know – note to Ofqual - this lack of conferring helps see what knowledge voids exist)
	• What questions do you have, what would you like to know? – Anything that perplexes anyone about the process?
	• What are the challenges to reliable or accurate marking do you imagine?
	    ○ What factors make it more difficult?
	    ○ What factors make it more easy?
	• How do you imagine Exam Boards try to ensure accuracy /reliability and replicability of marking – what do they build in (probe any knowledge of what actually goes on)
	• Is that likely to be effective in your view?
	• What is involved in examining? (probe to see how much they understand)
	• Do you know any examiners?
	• Have you ever examined?
	• What type of person do you think becomes an Examiner? *(note image of the type of person is a projective technique image can affect desire to examine and faith in results)*
	• Have you ever considered it yourself? –why and why not?
	• Expose propositions briefly based on survey responses e.g.;
	    ○ It helps me prepare my pupils better
	    ○ It helps my own pupils get better grades
	    ○ It is good for my career
	    ○ It shows you have grit and commitment
	    ○ it is inherently interesting
	    ○ It "ruins your life" for a month
	    ○ They don't pay enough for the work
	    ○ School doesn't support you
	    ○ Not compatible with home life/personal commitments
	    ○ Time of year
	• Expose a schematic of the marking and standardisation process
	• How do they feel now they have seen this – better or worse about quality of marking – more or less faith in results?
	    ○ What is most reassuring/impressive?
	    ○ What least?
	    ○ Again, any questions?
	• Also expose data about the type of person who becomes an examiner (experience, training and background) : probe effect on perceptions about the accuracy of marking
		25 mins

Guidance		
	• What information are you given about marking and where do you get it from?	15 Mins

	• How they feel about the guidance they are able to give pupils in mock marking or preparation for exams
	• What guidance is there?
	• What is helpful?
	• What is lacking?
	• Prompt on the following if they do not come up:
	    o Value of the mark scheme
	    o Exam board guidance or training they have ever undertaken
	    o Published guidance - is it clear, does it help them (compare subject to subject)
	• Any particularly difficult subjects or papers to mark internally or guide pupils on – why?
	• Any easier to mark – why, guidance? Type of question?
	• Anything that can be obviously improved in terms of things the Exam Boards can do for you as teachers?	

Review examples of guidance etc. helpful to them in school environment		
	• Refer back to at least one example per person they have thought of and brought with them in the pre task (mark schemes, online guidance etc.) What is good /bad about that example of guidance etc.? Why?	15 mins
Summary		

	• One most important thing the Exam Boards could do to make marking and examining more attractive
	    o To them
	    o To the school?
	• Or one most important thing they could do to improve their perception of reliable and accurate marking as teachers?
	• What information would they like from the Exam Boards
	• What have you heard today that that has most impressed you or changed your view on accuracy of marking – if anything?
		10 Mins

Appendix C – Information on the marking system presented to teachers at the research sessions

MARKING SYSTEM SUMMARY
(hybrid of various systems)

QUESTIONS DISCUSSED AND REVIEWED BY PRINCIPAL AND CHIEF EXAMINERS AND VARIOUS EXPERT 'REVISERS' **QUESTIONS APPROVED BY REVIEW BOARD** **DRAFT MARK SCHEME WRITTEN**	EXAM BOARDS RECRUIT AND ASSEMBLE EXAMINER TEAMS – DONE BY CENTRAL EXECUTIVE NOT PRINCIPAL EXAMINER - MANY MAY BE ELIGIBLE INCLUDING UNIVERSITY LECTURERS, SCHOOL TEACHERS ETC TEAM IS CHIEF EXAMINER PRINCIPAL EXAMINER (WITH SENIOR TEAM LEADERS IF LARGE PAPER) V TEAM LEADERS V ORDINARY OR 'EXPERT' EXAMINERS (EG 6 PER TEAM LEADER)	EXAMPLE SCRIPTS ARE CHOSEN FOR PRACTICE STANDARDISATION AND SEEDING PRINCIPAL EXAMINERS REVIEW AND AMEND MARK SCHEME IN LIGHT OF WHAT PUPILS HAVE ACTUALLY SUBMITTED MARK SCHEME CHANGED IN LINE WITH ANY ISSUES TEAM LEADERS ARRIVE AT 'PRE STANDARDISATION MEETING' AND ARE BRIEFED TEAM LEADERS MARK THE PRACTICE AND STANDARDISATION SCRIPTS AND ARE FULLY BRIEFED	STANDARDISATION OF ORDINARY EXAMINERS TAKES PLACE EXAMINERS COMPLETE PRACTICE SCRIPTS EITHER IN FACE TO FACE FORUM OR NOWADAYS MORE FREQUENTLY REMOTELY ON THEIR PC OR IN AN ONLINE ONLINE FORUM WITH TEAM LEADERS	EXAMINERS COMPLETE STANDARDISATION SCRIPTS (TYPICALLY 10) THEY ARE CONTACTED BY TEAM LEADERS TO DISCUSS IF THERE ARE ANY ISSUES THEY EITHER MARK ONLINE AND READ THE COMPUTER COMMENTS OR SEND SCRIPTS IN FOR MARKING AND COMMENT BY TEAM LEADERS EXAMINER IS EITHER CLEARED TO MARK OR ASKED TO DO MORE SCRIPTS THE TEAM LEADER BACK READS THESE AND FINALLY CLEARS THE EXAMINER TO MARK IF THEY QUALIFY

EXAM IS SAT

EXAMINERS MARK FOR VARYING PERIODS BETWEEN 2-4 WEEKS NUMBERS OF SCRIPTS ALSO VARY 350 SCRIPTS PER PERSON FOR GCSE UNITS OR 200 FOR A LEVEL

DURING THIS TIME THEY SUBMIT THEIR MARKS AND SCRIPTS FOR SCRUTINY AS THEY PROCEED

THERE ARE TWO SYSTEMS ONLINE MARKING OR WHITE PAPER/WHOLE SCRIPT /PEN AND PAPER MARKING

Two systems

Online   **Pen and paper**

ONLINE MARKING

ALL SCRIPTS ARE SCANNED IN AND THE EXAMINER MARKS ONLINE UPLOADING HIS SCRIPTS AND MARKS AS HE GOES .
THE BOARD INCLUDE ANONYMISED 'SEED' OR TEST SCRIPTS AS THEY GO TO TEST THE EXAMINER IS MARKING WELL AND ACCURATELY AND IF THE EXAMINER 'FAILS' ANY THEY ARE CONTACTED ANDTALKED THROUGH THE ISSUES
TEAM LEADERS CAN CHECK HOW FAST AN EXAMINER IS MARKING AND AT WHAT TIME OF DAY AND SAMPLE RANDOMLY AT ANY POINT

PEN AND PAPER MARKING

TEAM LEADERS RE VIEW SAMPLE SCRIPTS FROM ALL EXAMINERS IN TWO BATCHES OR SAMPLING POINTS

THEY CHECK THAT EXAMINERS ARE MARKING CONSISTENTLY AND WITHIN AN ERROR MARGIN

**IN BOTH SYSTEMS EXAMINERS CAN CALL TEAM LEADERS TO DISCUSS QUESTIONS ON MARK SCHEME AND HOW TO INTERPRET IT
OR PASS BACK SCRIPTS AND GET ANOTHER VIEW**

IF THERE ARE ANY ISSUES THEN SENIOR TEAM LEADRS OR PRINCIPAL EXAMINERS ALSO SAMPLE THAT EXAMINER'S SCRIPTS AND RE MARK

THEY CAN STOP AN EXAMINER AND RECALL THEIR PAPERS IF THEY ARE CONCERNED

OR

IN ONLINE MARKING THE SYSTEM CAN HALT AN EXAMINER UNTIL THE TEAM LEADER HAS DISCUSSED ISSUES WITH THEM

MARKING IS COMPLETED AND ALL SCRIPTS SENT IN TO THE BOARD

GRADE BOUNDARIES SET AND GRADES ALLOCATED SENIOR EXAMINING TEAM ATTEND THIS

PEN AND PAPER MARKING MARKING REVIEW:
(NOT UNIVERSAL)
CENTRES AND INDIVIDUAL CANDIDATES WHOSE MARKS ARE DIVERGENT BY TWO GRADES OR MORE FROM PREDICTED ARE RE-MARKED AND THE WORK OF EXAMINERS WHO SEEM ERRATIC IS REVIEWED

EXAMINERS RECEIVE AN ASSESSMENT GRADE FROM THE BOARD A FEW MONTHS AFTER THE PROCESS ENDS

We wish to make our publications widely accessible. Please contact us if you have any specific accessibility requirements.

First published by the Office of Qualifications and Examinations Regulation in 2014

Any enquiries regarding this publication should be sent to us at:

Office of Qualifications and Examinations Regulation

Spring Place	2nd Floor
Coventry Business Park	Glendinning House
Herald Avenue	6 Murray Street
Coventry CV5 6UB	Belfast BT1 6DN

Telephone 0300 303 3344
Textphone 0300 303 3345
Helpline   0300 303 3346