



England's 'examinations industry': deterioration and decay

**A report from HMC on endemic problems with
marking, awarding, re-marks and appeals at
GCSE and A level, 2007-12**

September 2012

Contents

Preface

Letter to the Secretary of State

SUMMARY

The nature of the problems

1

Overall diagnosis
Hidden problems – their nature and scope
What is the evidence and who suffers?
- *The school sample*
- *Methods of collecting data*
- *Who suffers?*
What is to blame?
What should be done?

Seven failings of England's 'examinations industry', 2007-2012: HMC's evidence

Unsatisfactory awarding of grades

8

Failing 1. Long-standing year-on-year variations in grades awarded in the same subject at the same level

Failing 2. Unexplained / unsatisfactory boundary changes to a previously stable subject

Failing 3. Significant and widespread variations between awarding bodies in the percentage of top grades awarded at A level across all subjects

Failing 4. Important and unexplained variations between some key subjects in the percentages of top grades awarded at A level across all boards

Poor quality marking

19

Failing 5. Persistent and widespread incompetence in marking

Failing 6. Erratic and inconsistent marking and re-marking on a large scale in relation to a specific examination

Obstructions to redress: re-marks and appeals

26

Failing 7. Significant doubts about the accuracy and fairness of the appeals procedure

Appendix: Historical evidence of poor examination procedures at a typical HMC school

31

Preface

The Times 19 September, 2012

The Education Secretary is to be commended for seeking to ensure that able children are challenged by public examinations. The flight to International GCSEs, caused by the failure of home-based GCSEs adequately to prepare pupils for advanced studies, may now be halted if he succeeds in his aims. However, he must do everything in his power to ensure that the foundations of our examination system will be as sound as the new baccalaureate will be rigorous. Every year poor and sometimes incompetent marking leads to unfairness and injustice for many students; and too often, when pupils and schools cry “foul”, those responsible for examining take refuge in the protective cover of an appeals system which places an adherence to process ahead of the quality of marking.

Any reforms to the character and content of the 16-plus examination “superstructure” will be seriously undermined by systemic weaknesses in the foundations of the system. Indeed, without careful reform of the assessment “industry” itself and the regulation of that industry, any reform, such as limiting competition between awarding bodies potentially making them less responsive to market forces, might actually make matters worse.

Christopher Ray,
Chairman, HMC

The Independent 19 September, 2012

There is a good deal to commend Michael Gove's plans for the reform of 16-plus examinations. Those of us who have deserted domestic GCSE and chosen, instead, to enter students for the international variant cannot doubt the attraction of greater rigour and challenge at this level.

However, he must ensure that the foundations of his proposed system are as rigorous as the content. There is a crisis in the systems for assessment that goes way beyond current concerns about English GCSE, and that will not be set right simply by franchising subjects to exam boards. Marking is routinely unreliable. Wide variations in standards and in grade allocations appear to be endemic. Mark schemes are ill-informed and limiting. The procedure for appeals is time-consuming, partial and opaque.

If Mr Gove's reforms are to be effective, he must undertake a thoroughgoing review of this essential element of the examinations industry. Without it, the English Baccalaureate will quickly become as discredited and devalued as its predecessors.

Kenneth Durham,
Vice-Chairman, HMC.

Daily Telegraph 20 September, 2012

For some time there has been wide consensus among school leaders in both independent and maintained schools that GCSEs require a significant overhaul. They have become narrowly focused, contain too much assessment and are a poor preparation for advanced study. Thus, the Coalition's plans for reform are broadly to be welcomed.

But this is only half the answer. The flight by independent schools to International GCSEs has been due also to schools' lack of confidence each year in marking and, when things go wrong within the exam boards, in obtaining justice for students who have been awarded the wrong grade.

An overhaul of the examinations industry, as well as qualifications, is what is required.

Dr William Richardson,
General Secretary, HMC

Rt Hon Michael Gove MP
Secretary of State
Department for Education
Sanctuary Buildings
Great Smith Street
Westminster
London SW1P 3BT

Leading
Independent
Schools



Dear Secretary of State,

Thank you for the interest you have shown in work conducted at HMC to pull together informed analysis of top priorities for the reform of examinations in England. We believe this work to be especially timely in the context of the preparations currently being made for qualifications reform. We hope that it will assist you and your officials.

We believe the changes envisaged to be overdue. The culture of GCSEs is now outmoded and there are clear improvements that can be made to A levels.

As we emphasise in our report, the biggest threat to the new qualifications that are planned is the widespread, seemingly random and largely unexplained inadequacy of examining. This is an urgent problem.

While there are many excellent examiners, including a large number in our own schools, we detail here just how widespread is the incidence of poor marking, the huge and distracting recourse made by schools to securing the re-marking of papers, and the highly unsatisfactory nature of an appeals system sustained by the legacy of a regulatory environment that has allowed examination boards to avoid accountability by hiding behind protocol.

We hope to continue working with Ofqual and with your officials in the identification of priorities and speedy steps that can be taken to arrest a growing disillusionment in examining that we depict in these pages.

We also welcome the support for our work received from ASCL. The issues discussed here and the often shocking data that has brought them into focus are pertinent to all schools and colleges in England.

Yours sincerely,

Dr Chris Ray
HMC Chairman

Dr William Richardson
General Secretary, HMC

SUMMARY

- HMC welcomes proposals to overhaul GCSEs and A levels. It welcomes especially moves to increase rigour in subject studies, reduce the burden of assessment on students aged 15-18 and differentiate student achievement more clearly across the grade range.
- In this report HMC sets out why these changes to qualifications (the superstructure) are almost certain to be undermined by long-standing failings in how young people are examined (the foundations). Unless examining is reformed substantially, the introduction of revised qualifications will amount to new houses built on existing sand.
- We welcome the interest that the Secretary of State has shown in our work and commend to him our findings. They are highly relevant to the implementation of qualifications reform.
- The report summarises what is known nationally about the confidence that current 'users' of GCSEs and A levels have in these examinations. Although we are accustomed in our schools each year to fighting repeated injustices in the marks and grades handed out by the exam boards in England, it was shocking to review national trends that point to a steady erosion of public and professional confidence. Across the country in 2011-12:
 - marking is widely seen as unreliable:
 - Those not confident in the accuracy of GCSE marking:
Parents 41% Teachers: 38% Students: 28%
 - Those not confident in the accuracy of A level marking:
Parents 45% Teachers: 27% Students 30%
 - the number of enquiries teachers make about the accuracy of results, the number of grades that are changed each year and the percentage of grades that are changed are all increasing;
 - almost half of all teachers (42%) say that they have to rely on the 'enquiries about results' procedure (i.e. a formal complaint) to secure accurate marks or grades for their students.
- Despite such figures HMC believes that understanding of the problems that schools and their students encounter is low, both in policy circles and in the public sphere more broadly.
- To assist this understanding we publish in our report key example of what goes wrong, how much is known about why (much remains unexplained due to a culture of secrecy in the exam boards and lack of focus in Ofqual) and the wider implications of each of these failings.
- Specifically we detail seven failings of the current 'examinations industry' in England (pp. 9-29), grouped under three headings:
 - Unsatisfactory awarding of grades
 - Poor quality marking
 - Obstructions to redress: re-marks and appealsEach of these failings is characterised by a precise definition of the problem concerned, our evidence and findings in relation to it (including what is known, surmised or unexplained), followed by an assessment of the wider implication of these findings. Local examples are provided which bring each problem vividly to life.
- The authority for our findings derives from several sources: national data; collaborative work with schools and subject associations in the maintained sector; internal HMC surveys; and data from groups of HMC schools, particularly from heads of departments. In national terms the staff in our schools are exceptionally well qualified in subject knowledge and our schools are part of an independent sector that government research shows to be the most expert in the country at predicting student grades accurately. In important respects, our findings about erratic and unexplained examining are strengthened by the high level of classroom continuity characteristic of our schools.

- The oldest information dates from 2007 (and is in the Appendix and the section on appeals) but most data are drawn from systematic monitoring conducted over the last three years (2010-12). More than half of the data in our report have been collected within the last month.
- Our findings readily suggest areas for immediate and medium term action. These are brought together on p.6 ('What should be done?'). Many are tasks for Ofqual. Not a few are urgent and there are several to which we and others can make a highly informed contribution.
- In essence, the challenge in front of Ofqual is to scrutinise much more closely what goes on inside the exam boards, especially concerning marking which is the Achilles heel of the entire 'examinations industry'. It also has work to do in transforming the effectiveness of an outdated appeals system redolent of another, altogether cosier age.
- We welcome recognition by ASCL of the significance of our enquiry. We have published this report on behalf of all students in England who do not receive on results day marks or grades that accurately reflect their performance and achievement. As Brian Lightman, General Secretary of ASCL says, our findings are important and are '*likely to have uncovered the tip of an iceberg*'.

The nature of the problems

Overall diagnosis

As indicated in the preface, HMC welcomes in broad terms current initiatives to overhaul GCSE and A level qualifications in England. In particular, we welcome a focus on increased subject rigour, reduction in the volume of summative assessment within each qualification and the aspiration to use public examinations to differentiate more clearly and across a wider spectrum the breadth and depth of understanding achieved by students.

However, we characterise these plans as houses that would be built on the sand of a deteriorating national industry of public examinations. HMC schools and their counterparts in the maintained sector experience the unstable foundations of national assessment every year. For our part, we are certain that this problem will continue and thus erode confidence in new qualifications unless clear steps are taken to strengthen the basis in examining on which they are constructed.

Many HMC schools have already voted with their feet. International GCSEs have taken off in spectacular fashion as disillusionment with GCSEs has grown. Sparked in the mid-2000s by a desire to move away from compulsory coursework, this surge is now associated with wider dissatisfaction over examination and grading. This year, IGCSE entries from independent schools increased by 52% and now accounts for one quarter of their GCSE/IGCSE entries combined.¹ Meanwhile, at 16+ our schools have been diversifying into other awards beyond the domestic staple of A levels – the IB Middle Years and Diploma programmes and the Pre-U examination.

This report is in two parts. First, we review the national evidence of low professional and public confidence in public examinations and the growing incidence of the marks and grades initially given to students having to be changed. In the second part we illuminate this national picture by throwing direct light on eight specific shortcomings in examinations and their regulation in recent years. We believe that schools and colleges throughout England will recognise and have experienced similar failings. In particular, we present our evidence on behalf of the large number of students annually across England who do not receive on results day marks and grades that accurately reflect their performance and achievements.

Hidden problems – their nature and scope

Every year many thousands of students in England are awarded the wrong grade in one or more of their GCSE or A level examinations. And every year schools and colleges are faced with diverting resources into correcting this problem. This arises

¹ *ISC Year 11 Exam Results 2012*. Press Release: Independent Schools Council, 1 September, 2012. The increase in recent years (% of all Year 11 entries comprising IGCSE from schools reporting data) is: 2007, 4%; 2010, 11%; 2011, 17% (source: ISC).

in the main from a fact that all of the main examination boards concede in private: the quality of marking is not good enough.

Ofqual statistics and commissioned surveys outline the national trend in stark terms. These show that the number of enquiries teachers make about the accuracy of results, the number of grades that are changed each year and the percentage of grades that are changed are all increasing.² One in 40 of all entries are subject to a teacher enquiry about marking accuracy and almost half of all teachers (42%) say that they have to rely on the formal 'enquiries about results' procedure to get accurate marks and grades results for their students.³

Every so often – in 2002 over A level grades leading to university admission or in 2012 over GCSE English grades leading to education, training or employment post-16 – a *cause célèbre* hits the headlines and occasions significant political activity.

This report details the less newsworthy but widespread problems that students and their schools routinely face in attempting to secure correct grades, across a range of subjects and at a range of ages and levels.

The report attributes the main cause of this to poor marking. But equally important, students and their schools have to fight hard to secure corrections at subsequent stages: the re-marking of written work, review of the level at which grade thresholds in any one year have been set ('awarding') and the formal appeals process that is invoked whenever schools are convinced that re-marking or review of awarding has still not corrected a glaring problem.

Put briefly, we are certain that the culture of examining in England is not working well enough in its fundamentals or in its detail. Fundamentally, there is far too much examining – and re-examining – of students as they proceed through each subject course. In terms of detailed operations, the boards are still too secretive and the manner in which the response to complaints is regulated is too concerned with impersonal procedures and not enough with whether scripts have been marked accurately.

For each student and for their school or college the stakes are high. But the 'system' prioritises aggregate results that are impersonal yet conform to statistical models of how the cohort of candidates as a whole may be expected to perform. GCSE English in 2012 is a case in point. At the time of writing the regulator is assuring the

² *Statistical Bulletin Appeals against Results for GCSE and GCE: Summer 2011 Examination Series*, Ofqual: Coventry, March 2012. According to this Bulletin, in 2011;

- 12,250 grades changed at GCE (0.49% of subject grade awards) and 26,270 grades changed at GCSE (0.45% of subject grade awards) as a result of schools and colleges following the 'enquiries about results' procedure.
- Moreover, there number of enquiries about results was in 2011 was 204,600. This comprised 2.47% of all subject awards, that is 1 in 40 of all subject entries a rise from 171,400 in 2010 and resulted in 24,200 (0.40%) changes of subject grade at GCSE and 10,500 (0.47%) changes of subject grade at GCE.

The 204,600 enquires about results in 2011 represents 2.47% of all subject awards, that is 1 in 40 of all subject entries.

³ *Perceptions of A levels, GCSEs and other qualifications: Wave 10*. Ofqual: Coventry, March 2012, p. 45.

public that national standards have been secured, based on setting grade boundaries in relation to tests that 16 year olds sat when they were age 10 or 11. Little wonder, then, that it concedes that it cannot account for results reported from schools that show high variance of grades (in either direction) compared to previous years.

It should not be the job of schools, or groups of schools, to have to estimate the extent to which the marks awarded in public examinations are accurate and reliable. Our fundamental task is teaching and the nurture of the young people in our charge. Neither are we set up or resourced to shadow comprehensively the work expected of the examination boards and the regulator. However, each year we get sucked onto this territory because we are not confident that large numbers of our students are getting their just desserts. And if we are not confident, neither are our students or their parents.

What is the evidence and who suffers?

The school sample. HMC represents the heads of 221 schools in England and a further 31 in the remainder of the UK many of whose students sit English examinations. The staff in our schools are very highly qualified in terms of subject knowledge.⁴ Our schools are part of a sector of independent schools which, compared to all other types of school or college in England, are by some distance the most accurate in predicting the grades that students will secure in public examinations.⁵ They also have to put aside significant resources each year to tackle the problems in marking, awarding and appeals that repeatedly occur.

Our schools have neither the time nor should they have a remit to prove on a statistically significant scale the precise extent to which things go wrong. We do, however, know our students very well and have an unrivalled track record in assessing their performance realistically and accurately. As such, our schools comprise a very good sample to act as 'early warning' of wider system failings. In turn, this gives us a strong level of confidence to point to evidence of problems highly likely to exist on a large scale nationally.

For this reason Brian Lightman, General Secretary of ASCL, has strongly endorsed the impulse behind our report.

'HMC has been sharing with us for some time its concerns and mounting evidence base of aspects of assessment and examining that are not working well. This work of our colleagues in HMC schools is important and is likely to have uncovered the tip of an iceberg'.

Maintained schools and FE colleges face the same problems with getting fairness for pupils from the examination system. However, these schools and colleges do not

⁴ ISC Teacher Survey. *ISC Bulletin* 26 (May, 2012), pp. 8-14.

⁵ BIS Research Paper 37 *Investigating the Accuracy of Predicted A Level Grades as part of the 2009 UCAS Admissions Process*, Department for Business Enterprise and Skills, London. June 2011, pp. 8, 25-26, 29

always have the resources to pursue complaints and appeals with the vigour of their independent school counterparts.

Methods of collecting data. Each year HMC asks its members to report problems with assessment and examinations. Each year most schools have problems in at least one area and, with varying degrees of success, attempt to right obvious failings in examining through the various processes for redress that exist. Often the problem turns on the failings of an individual marker that distort local results.

Each year also one or two problems emerge, sometimes quite quickly but often more slowly after results day which point concerns on a scale indicative of a systemic failing beyond the vagaries of a 'rogue marker'. Either way, the impact is equally severe on the student who receives the wrong mark for a module or the wrong grade for an entire course.

In this report, we detail both kinds of problem – how 'rogue-marking' can take a significant amount of time and energy to be conceded by the examination board concerned and how patterns of evidence can emerge which tell us that something is wrong on a larger, more 'systemic' scale. We also show how when a school is certain that injustices have occurred the procedure for formal appeals is focussed on the wrong target. Instead of a chain of top-down instructions to establish that due process has been followed inside the boards, the spotlight should have been shone primarily on the evidence that initial marking was accurate and re-marking (as opposed to a review of the marks awarded) was undertaken.

The oldest information dates from 2007 (and is in the Appendix and the section on appeals) but most data are drawn from systematic monitoring conducted over the last three years (2010-12). More than half of the data in our report have been collected within the last month.

Who suffers? Nobody in the examination boards or at the regulator is inconvenienced if the national results for a particular exam conform to statistical models of predicted outcomes *and* if schools are too busy or have insufficient resources to pursue the often protracted and distracting task of securing proper review by the exam boards.

Meanwhile students from all kinds of school and college suffer if their teachers are convinced that they have been in receipt of the wrong mark or grade. Thus, it is of great concern to discover that:

- 'one in five of all teachers (20%) believe that around a quarter of GCSE students get the wrong grade';
- this can be described by Ofqual's contract researchers as a 'broadly positive' statistic;⁶ and
- 'the most frequently mentioned reason for A level teachers believing that the accuracy and quality of A level marking has decreased over the past two years is the lack of high quality or experienced markers'.⁷

⁶ *Perceptions of A levels, GCSEs and other qualifications: Wave 10.* Ofqual: Coventry, March 2012, p. 39.

⁷ *Ibid.*, p. 27

Each year many students suffer the effects of such administrative, procedural and educational shortcomings. HMC is now providing evidence from its surveys of members' schools over recent years to illuminate more clearly the nature of these failings.

What is to blame?

The recent hearing of the House of Commons Education Committee on the administration of examinations⁸ heard many opinions about where the problems lie.

Among these opinions were claims that in order to cope with the receipt of 15 million scripts each summer, and smarting from the accusation that for too long they were a 'cottage industry', the exam boards have become another sort of industry – one based on high unit volume, narrow 'specifications' (descriptions of the elements of a course), competitively positioned textbooks designed to secure market share among candidates, an overriding concern for reliability in assessment assured through on-line marking and a cumbersome, opaque appeals process.

Debates on these questions are now fully enjoined. Our report focusses on a relatively under-explored aspect but central aspect of professional and public confidence – reliability of marking and 'awarding', and adequate redress through the review and appeals process.

We have spoken of an 'examinations industry'. Restructuring of the English boards now looks certain but the authorities will be mistaken if they believe that, in itself, this will secure the foundations for the success of reformed GCSEs and A levels. These new qualifications will be just as vulnerable to the erosion of public confidence if the processes detailed in the following pages are neglected any longer.

The most recent survey commissioned by Ofqual (conducted November-December 2011)⁹ is not at all encouraging in terms of the confidence of teachers, students and parents in the accuracy of marking and awarding at GCSE and A level.

GCSE

- Those not confident in the accuracy of GCSE marking:
Teachers: 38% Students: 28% Parents 41%
- Those not confident that *most* students get the right GCSE grade
Teachers: 23% General public (parents and students) 33%

A Level

- Those not confident in the accuracy of A level marking:
Teachers: 27% Students 30% Parents 45%
- Those not confident that *most* students get the right A level grade
Teachers: 13% Students: 24% Parents 33%

⁸ House of Commons Education Committee. *The Administration of Examinations for 15-19 year olds in England*. 3 Vols. [HC 141 I-III]. .

⁹ *Perceptions of A levels and GCSEs – Wave 10*. Ofqual: Coventry, March 2012, pp. 23-24, 27-28, 35-36, 39.

Each year individual students are given the wrong marks and grades, for reasons that vary annually, are unpredictable, too often remain unexplained and undermines significantly the confidence of teachers in all schools and colleges. And each year redress remains costly and difficult for hard-pressed schools to secure.

What should be done?

The steps leading to improvement include short-term and longer-term measures.

Immediately. There are some very practical steps that Ofqual can take, working with informed partners and with the endorsement of the Department for Education, to tackle some of the features of examinations currently most corrosive to public and professional confidence.

This will involve a return by Ofqual once more to a reconsideration of one of its abiding principles: 'use of statistics to help inform examiner judgements'. As this report amply demonstrates, the perspective of parents, students and teachers in HMC schools is that, too often, deterministic statistical models designed to secure national standards appear to trump regulatory scrutiny of the accuracy and fairness of marking (and re-marking) at the local level.

Our schools are not alone in being unconvinced. As reported above, 42% of teachers say that they have to rely on the 'enquiries about results' procedure to secure accurate marks or grades for their students. Unconvincing marks and grades should not be leaving the offices of the exam boards on anything like this scale.

Moreover, markers are currently examining relatively simplistic modules. How will they and the boards cope once synoptic questioning in end-of-course examinations becomes much more common again? At that point how will the use of national statistics support a new emphasis on local examiner judgement? And if 'drift' in grade boundaries is to be patrolled from outside the boards much more actively in the future, how will students with Specific Learning Difficulties be safeguarded from unwitting indirect discrimination, should new kinds of examining challenge current practice in learner support and access arrangements?

Prompted by these questions, the data in our report suggest starting points for these aspects of Ofqual's work. Our suggestions are that it should:

- investigate who the boards' markers are. What are their qualifications? At what stage are they in their careers?
- publish the data on markers' performance that the exam boards are obliged to maintain. Thus, it should require the exam boards to inform the regulator, and make public, the percentage of examiners who are graded unsatisfactory and satisfactory in any one year;
- scrutinise its own evidence base and report on the patterns of unsatisfactory examining that this might reveal;
- report estimates of the real incidence of marker unreliability (given that schools tend not to appeal when marks unexpectedly increase);
- explain why the percentage of top grades varies between boards across all subjects;

- track and account for the wide variations in results in the same subject, year on year, of the kind we outline below. Why do these occur?
- review whether the emphasis in the current Ofqual Code of Practice in respect of re-marks (and re-moderation) can continue to be justified. Public (especially parental) and professional confidence in the appeals process would be significantly strengthened were the boards required to undertake *ab-initio* re-marking (rather than merely procedural 'review' of marks already determined);
- in view of the widespread occurrence of erratic marking that we identify, move to review the conditions which trigger the re-marking of an entire group of candidates for an examination from an individual school.

Medium term. Beyond these important operational improvements – all of which need specific attention – there are other more fundamental steps to be taken. For example, there is a need to start straightaway on addressing challenging tasks such as undertaking a serious review of how subjects are examined and the volume of assessment they require.

The context for this should be a staged reduction in the amount of assessment that GCSE and A level candidates experience (as is now planned).

Accompanying this, schools will want to see:

- clear evidence of increased openness and responsiveness on the part of the examination boards in respect of complaints, supervised and monitored by the regulator;
- progress toward the goal of both less assessment and better assessment of students;
- serious consideration being given to reducing the amount of external assessment at age 16, other than in the core subjects of maths, English and science, with a view to replacing it by school-based assessment verified by trained expert in-school assessors; and
- Ofqual charged with requiring exam boards to devise stimulating and challenging examination specifications which captures the interest of students and teachers and motivates them.

The core point here links to professional and public perception of the examining process and thus confidence in it. Most people, including many schools, think re-marking and re-moderation are *ab initio* processes. They are not; they are simply a review of the original marking to ensure that the agreed mark scheme has been applied correctly. There is a strong argument therefore for suggesting that the Code should be revised to require exam boards to carry out an *ab initio* re-mark (re-moderation), otherwise it will always appear that it is a mere rubber stamping exercise by the exam board.

Seven failings of England's 'examinations industry', 2007-2012: HMC's evidence

Unsatisfactory awarding of grades

The first four failings summarised from our evidence relate primarily to aspects of the way in which groups of students are awarded summative grades at the end of their course study.

The key process here is 'awarding' (the setting of grade boundaries) but, as the examples that follow demonstrate, the grades given to groups of students can also be affected by erratic and unreliable marking, or by inconsistent marking when entire subjects are compared.

Failing 1

LONG-STANDING, YEAR-ON-YEAR VARIATIONS IN GRADES AWARDED

What is the problem?

Schools receive unexplained and very large variations in the percentages of grades given to successive annual cohorts of pupils from the same school in the same GCSE or A level subject. This is despite the subject being taught by stable teams of experienced staff to annual groups of students of largely similar ability.

Headline findings

- Between 2010 and 2011, HMC uncovered wild variations in many schools in the grades awarded in English and English Literature GCSE across all exam boards; these fluctuations have amplified longer-standing concerns.
- Data on this was communicated to Ofqual on request in September 2011 and can now be seen as an 'early warning' of the problems that in 2012 have escalated into a national grading crisis.
- In 2012 many schools are concerned yet again about English GCSE grading compared to previous years and to forecasts.

Wider implications

From HMC data it seems very likely that for a number of years tens of thousands of students across schools in England have been receiving questionable GCSE grades in English, English Language and English Literature.

What specifically has gone wrong and what is the evidence?

Between 2010 and 2011, variations of over 10% occurred in the award of GCSE A* or A*/A grades for both English and English Literature in at least one in five of those of our schools that enter candidates for GCSEs, but probably more.

The average number of candidates entered for English in each of the 42 schools concerned was 114 with a total of 4,769 candidates being affected. The figures for English Literature were similar. Most schools (32) reported variances in either direction of 15% or more – a figure far higher than can be explained by differences in the ability of successive student groups or changes in staffing. None of the 42 schools had changed exam board that year.

For schools with 100+ entrants, the largest reported variances between the two year groups at either grade or in either subject were: 32%, 34%, 35%, 44% and 67%.

Problems were encountered with each of the main exam boards. Variances of 10%+ absolute are a serious concern. Variances of 20% are completely unacceptable and their causes require urgent attention.

So far in 2012 we have asked schools to alert us only in general terms if they have concerns in the grading of GCSE English and English Language GCSE. By 19 September, 59 had done so (amounting to more than a quarter of all HMC schools offering GCSE). 24 report concerns over the setting of grade boundaries at A*/A and A/B. 25 report concerns over the setting of grade boundaries across the range A*-C/D. Ten report concerns over the setting of grade boundaries specifically at the C/D boundary.

Local example

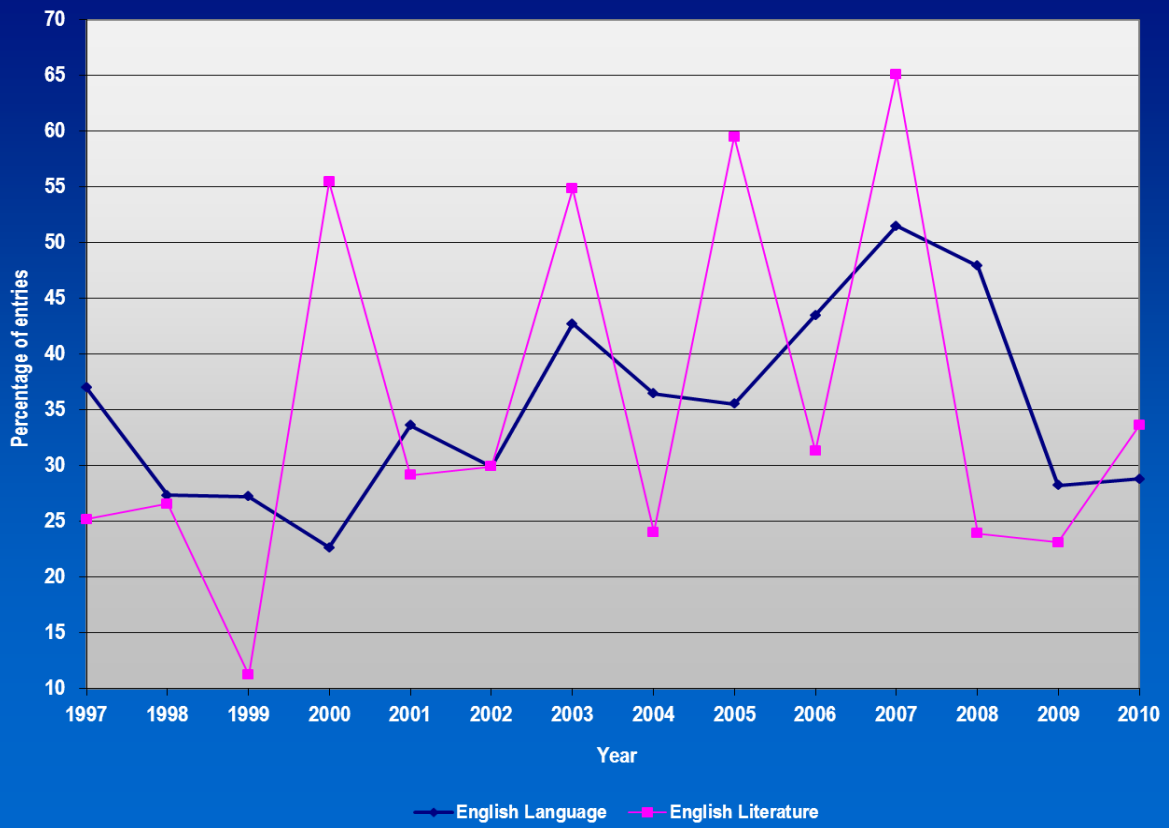
A striking example of the kind of wild fluctuations that schools can experience over more than a decade are those recorded at King Edward's School, Birmingham.

The English department at King Edward's has been stable for many years, with an unchanged head of department and a very stable contingent of pupils year-on-year in terms both of size – 120, plus or minus five – and ability.

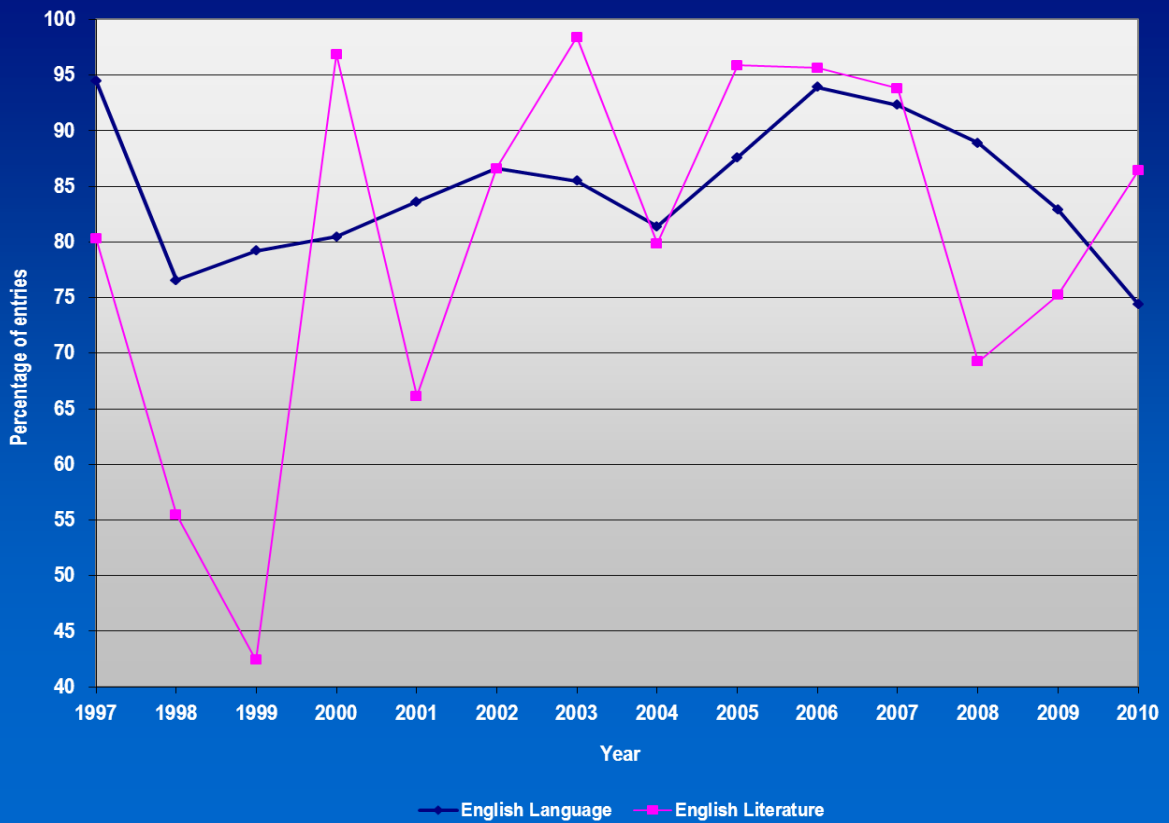
The apparent randomness of the school's GCSE English Language and English Literature results at grades A* and A is illustrated in the attached graphs. A* proportions ranged from 11% to 65% and back again during 1997 to 2010; A*/A over the same period ranged from 43% to 98%. Since the school moved to IGCSE two years ago, for the first time it has had results consistent from year to year and across the two subjects.

Chief Master John Cloughton commented: *'we had a row with the examiners about five years ago and got absolutely nowhere. As a sop they sent the Chief Examiner along to help our department and the next year the results got worse. In the end, we simply and moved'*.

GCSE English Language/Literature Comparison at A* 1997-2010



GCSE English Language/Literature Comparison at A*-A 1997-2010



Failing 2

UNEXPLAINED / UNSATISFACTORY BOUNDARY CHANGES TO A PREVIOUSLY STABLE SUBJECT

What is the problem?

A boundary change in a GCSE or A level award unexpectedly, cannot satisfactorily be explained, is not corrected and appears to result from faulty marking.

Headline findings

- After years of stable grades, HMC (and other) schools have experienced dramatic changes in grades awarded in GCSE AQA Drama.
- Schools have identified excessively harsh marking of one unit but the board denies the existence of significant problems and continues to justify the large majority of grades awarded initially.
- Despite the board initiating a review, it never became clear whether the board had refused to give ground for a hidden reason – because the performance of all candidates (i.e. the grades they receive) have to be matched to national statistics in predicted grades across all schools.
- This lack of transparency is disillusioning for teachers and misleads students into questioning their abilities.

Wider implications

Schools are left dissatisfied and unconvinced. Many have moved to another board in the hope of obtaining greater transparency. The dissatisfaction of the schools undermines student confidence that each has received a fair grade. In the light of this failing, government proposals to award the examining contract for each GCSE subject to an individual board, thus eliminating school choice, could make matters worse if exam board culture remains unchanged.

What specifically has gone wrong and what is the evidence?

A specific problem in relation to GCSE Drama awards emerged on a large scale in 2011 following more isolated concerns in previous years. Problems were tracked in detail in 2011, with HMC helping to co-ordinate evidence-gathering with state maintained schools. As a result, over 100 schools lodged complaints with the AQA board. The board set up a review involving many of these schools.

One of the schools concerned has a record over the previous 12 years of 92-100% A*/A grades, yet in 2011 achieved only 18% A*/A grades. Board representatives responded by saying that awarding in 2010 and 2011 had been 'extremely similar'.

The exercise exposed the existence of a number of incompetent examiners, one of whom had been graded excellent by the board during the course of the marking process. However, the senior examiners involved steadfastly refused to accept any culpability and, backed by the board, confirmed most of their decisions about the grades to be awarded. This was despite the board acknowledging that there had been 'some marking issues with the application of the mark scheme', thus hinting

that problems had stemmed from the adequacy of the question paper. A candidate at one school had received a re-remark increase of 37 raw marks out of 80.

Many schools in HMC and elsewhere have now deserted AQA Drama, citing loss of confidence in the board. Of those remaining, many have notified us of further problems in awards made in 2012. This year common complaints are that the exemplar material provided by the board to teachers is again inaccurate, the marking is again erratic and that internal board rules have made it harder for schools dissatisfied with marks received initially to trigger a group re-mark (see also *Failing 7*, below).

Local example

Eastbourne College is one of the many schools to have been left completely dissatisfied with this situation.

The drama department at Eastbourne was closely involved in challenging AQA on the marking of its written paper at GCSE 2011. It had with other schools an overwhelming concern at the standard and clarity of marking. It considers this to be wildly inconsistent and negative, and to have been accompanied by an unjustified and unexplained raising by the board of the grade boundaries across all grades.

One Eastbourne College student achieved 9 A* grades at GCSE but, completely unexpectedly, secured only a grade B in Drama and a grade D in the written paper element.

Re-marking has also been erratic. Four papers have been received by Eastbourne, re-marked from this summer's (2012) examinations – two completely unchanged, one up by 19 marks and another by 21 marks.

After the major problems in 2011 this summer, once again, dozens of schools have contributed to email discussion complaining of inconsistent and unfair marking, that promises made by the board have not been honoured and that, once again, the exemplar material has been inaccurate making it impossible for teachers to guide students accurately.

Headmaster Simon Davies said: *'This year, as last, we have been left very unsatisfied with the service provided by AQA. It is a small comfort to learn from numerous emails circulating widely among heads of school drama departments that our experiences and complaints are widely shared.'*

More schools in 2012 have made the decision to change boards. One head of such a department wrote to her peers in other schools this August: *'We lost our stage 2 appeal last year, regardless of another watertight case... Good luck to those of you embarking on another AQA battle again this year and to those of you changing boards. I am only sorry that they haven't sorted themselves out'*.

Another head of department wrote:

'For three years from 2009, we had a truly wretched experience with AQA GCSE Drama including a Stage 3 Appeal where we had a watertight case and it was clear

that everyone in the room understood that the grades awarded to our pupils were completely unjust - yet the appeal was not upheld. [Having now moved board] it is still very painful to us that for three years so many of our pupils were denied the grades that their talent and industry deserved - and there is nothing that we can do to change that now'.

Failing 3

SIGNIFICANT AND UNEXPLAINED VARIATIONS BETWEEN AWARDING BODIES IN THE PERCENTAGES OF TOP GRADES AWARDED AT A LEVEL ACROSS ALL SUBJECTS

What is the problem?

It is not clear why the percentage of A*, A and B grades awarded board-by-board across all subjects should vary to the extent they do. Ofqual and the boards are supposed to guarantee inter-board standards.

If patterns of this kind derive from the overall 'ability' profiles of candidates entered for the respective boards, the boards have yet to demonstrate this publicly.

Headline findings

- HMC surveys cannot explain why the three English exam boards which have the largest number of entries from our schools vary so widely each year in the proportion of top grades they award, when set in the context of an overall candidate group from our schools broadly comparable in ability and preparation by their teachers.
- The variances evident in HMC's data also exist in national figures for candidates from all schools across all subjects.
- HMC has engaged Ofqual on this issue but, as yet, we are no nearer a clear explanation of the causes of such disparity.

Wider implications

Variations of this kind do not inspire confidence in exam board comparability. This uncertainty is not helped by apparent indifference on the part of the regulator. Perhaps inter-board variation is explained with reference to differing profiles of ability across all entrants. If so, it is for awarding bodies to set out this evidence. If there is a subject effect (i.e. the subject mix offered by each board attracts student groups of different ability) this, too, needs to be documented publicly.

Other explanations that we would expect Ofqual also to have been commenting on include the possibility that some boards set easier exams than others or that, collectively, the boards mark scripts with varying degrees of rigour, or that some are more lenient in setting grade boundaries than others.

What specifically has gone wrong and what is the evidence?

Since the A* grade was introduced at A level in 2010, HMC has been tracking the attainment of its students at this level, board-by-board.

Our 2012 survey covered data from 125 schools that, between them, accounted for 42,841 candidate entries. Among boards to which over 10,000 entries were made in total by these schools, there was significant variation in the proportions of the highest grades awarded:

| | | | |
|--------|----------|------------|------------------|
| AQA | 14.0% A* | 48.4% A*/A | (13,749 entries) |
| EDEXEL | 23.3% A* | 60.7% A*/A | (10,635 entries) |
| OCR | 23.0% A* | 57.9% A*/A | (14,172 entries) |

Patterns in our surveys for 2010 and 2011 were very similar.

HMC is confident that these data are illuminating, comprising as they do, a relatively large sample of entries made by candidates of relatively similar ability.

Meanwhile, data for candidates from all schools and colleges in England are in the public domain. These also show significant variance among the boards in terms of proportions of the highest grades awarded (A*, A and B).

When HMC raised this matter with Ofqual in 2010, the regulator did not appear especially interested, beyond observing that it would be better if such variations as those in the HMC data did not occur.

Failing 4

IMPORTANT AND UNEXPLAINED VARIATIONS BETWEEN SOME KEY SUBJECTS IN THE PERCENTAGES OF TOP GRADES AWARDED AT A LEVEL ACROSS ALL BOARDS

What is the problem?

Between them, the boards and the regulator have done little to clarify why the proportions of high grades awarded by all boards varies in important respects, subject-by-subject.

Assuming the standards across boards are constant (a generous assumption given the data discussed under *Failing 3*, above) why do top grades vary nationally by subject? Are some subjects harder than others? Do some subjects attract less able students? And how can we explain why some subjects, commonly perceived to be 'hard' and known to attract very high attaining students, appear to be marked particularly harshly?

Headline findings

- In the national statistics, there are significant variances subject-by-subject in the proportion of top grades awarded (A*, A and B).
- These national data reinforce the judgement of heads of department in HMC schools that some subjects appear to be being especially harshly marked, all other factors being equal.
- While parity of marking and grade-setting between many subjects remains concerning but unexplored, detailed work has been done on modern foreign languages (MFL):
 - staff in HMC schools have been closely involved in the work of modern language associations (independent and maintained schools) to try to understand the paucity of top grades in modern foreign languages (MFL) when compared to other conventionally 'hard' subjects such as maths and physics;
 - research by these associations has shown that:
 - students taking a language in Year 12 (AS level) are more likely to drop the language than their other subjects;
 - compared to other students, those taking a language to A2 have the 'greatest discrepancy between the grading for ... that subject and their average score in the other subjects studied';¹⁰
 - having had these problems brought to its attention by the Independent Schools Modern Language Association (ISMLA) in November 2011, Ofqual is investigating them.

¹⁰ Association for Language Learning. *Severe grading at A level. Summary.* : available at http://www.all-london.org.uk/AL_ongoing_summary.htm.

Wider implications

If some subjects appear to be particularly harshly graded students will not embark on them. For those who do follow these courses, such apparent bias in grading is seen as unfair.

In recent years this problem has been at its most acute in modern foreign languages, nationally among our strategically important and most vulnerable subjects. 'Severe grading' seems to explain high MFL dropout at the end of Year 12 (AS level)

The inexorable reduction of entries in modern foreign languages over the last decade is well known and is related to perceptions of excessive difficulty, unreliable marking and harsh grading compared to other 'hard' subjects.

What specifically has gone wrong and what is the evidence?

The case study of modern foreign languages may provide a partial answer to the questions posed at the top of this section. Data reproduced below is taken from studies undertaken by ISMLA and the London branch of the Association of Language Learning (ALL). It should be read in conjunction with the fact that schools in ISMLA are those nationally that set the highest standards in predicting students' grades accurately.¹¹

The case study

- In 2009 students were first assessed at AS level against new courses in MFL. One school speaking for others commented: '*grade boundaries ridiculously punitive. 92/111 for A on French paper 1*' (these figures are significant – see last bullet point, below, and conclusion).
- 2010 was the first year that the new AS/A2 course could be completed in full and the A* grade could be awarded. Responses from around 50% of the ISMLA membership (80-90 schools; 3,000 candidates) found that just under half were disappointed with the numbers achieving the new top grade (i.e. in relation to their judgment about what it was reasonable for their students to have achieved).
- In 2010 also, marking of oral exams was '*widely viewed as unreliable ... [with] ... linguists known to be gifted being outperformed by weaker ones*'.
- In 2011 an ISMLA survey found 40% of schools disappointed with MFL AS grades across the grade range and 37% with A2 grades across the range. 50% of schools reported fewer pupils gaining the A* grade than predicted.
- In 2001 also, analysis by ALL suggested that compared to other subjects '*many linguists are good enough to cross the line for an A, but the majority of them are ruled out of the A* by the very high demands across our peculiarly multi-skilled set of examinations*'.
- In December 2011 Ofqual agreed to investigate, met with ISMLA representatives and subsequently wrote to say: '*we discussed with you the evidence we have from our modelling work carried out on the previous specifications that showed that grade A candidates in modern foreign languages tended to score a greater*

¹¹ BIS Research Paper 37 *Investigating the Accuracy of Predicted A Level Grades as part of the 2009 UCAS Admissions Process*, Department for Business Enterprise and Skills, London. June 2011, pp. 8, 25-26, 29

proportion of their marks as AS than grade A candidates in other subjects. If this has continued into the new [2008 onwards] specifications, it could go some way to explain why the proportion of A grades in modern foreign languages appears to be slightly lower than other subjects with a similar proportion of candidates at grade A¹.

Conclusion

When the new modern foreign language AS and A2 courses were introduced during 2008-10, it would appear that the boards placed at the wrong level in the mark-range the 'judgment grade' of A (all grade boundaries in A Levels are set in relation to the standards agreed by senior examiner judgement for grades A and E, as required by Ofqual). In addition, this has been compounded subsequently for A* awards by:

- (a) the 50/50 AS/A2 weighting in A level awards which makes it more difficult in MFL compared to other subjects to 'make up' ground in the A2 year; and
- (b) unreliability of oral examining, erratic marking of other papers and marks widely at variance with the school's rank ordering of candidates (all widely reported in HMC schools and all of which in December 2011 Ofqual undertook to investigate).

Local examples

The University of Oxford has declared at open days in 2012 that it will continue to require A grades rather than A* grades in languages. This is because the university recognises the special plight of A level language students and considers that those who secure 85% (UMS) have achieved enough to gain entry, despite not achieving an A*.

Other highly competitive universities set relatively low entry criteria (AAB-BBB), probably for similar reasons.

ISMLA Chairman Nick Mair of Dulwich College said recently: *'Many linguists in independent schools and elsewhere have given up any hope of seeing a correlation between linguistic competence and MFL examination results'*.

Poor quality marking

The nature of our evidence now moves to the most pervasive, corrosive and least well-documented aspect of the malaise endemic across English examinations – that of poor marking.

It is a problem conceded by each of the exam boards in private but one over which our schools find that they have to do battle in public on a persistent and, all too often, disillusioning basis. It is also a key failing in the conduct and administration of examinations that is far too little recognised and monitored by Ofqual.

Failing 5

PERSISTENT AND WIDESPREAD INCOMPETENCE IN MARKING

What is the problem?

HMC schools regularly and routinely have to challenge the exam boards about the performance of ‘aberrant’ or ‘rogue’ markers. Each year schools receive module scores at all levels (GCSE, AS and A2) and usually across several subjects that are, quite literally, unbelievable.

Headline findings

- The evidence is very widespread and, by its nature, hard to summarise other than to say that merely among HMC’s 250 schools, dozens of cases arise each year, every case has to be assessed within the school and judgments made about scale and seriousness, and the minimum redress to be sought.
- The case study set out below is not at all untypical in the extent of the problems identified by individual HMC schools over the last few years and the scale of the actions a school would deem necessary.
- Such is the scale of poor marking that, as Ofqual surveys show, teachers expect to have to mount challenges to it as a matter of routine.

Wider implications

More than anything it is this erratic and inconsistent marking, and its pervasiveness, that has eroded professional and public confidence in the reliability of the grades that students receive.

Most HMC schools are able to put the requisite resources behind following up the most blatant cases of poor marking (indeed, it has become a repeat element of budgeting). The concern of ASCL relates to the fact that the evidence from our schools almost certainly points to widespread injustices faced by all schools, many of which have fewer resources to divert to this work than do we in HMC.

The current political storm concerning GCSE English awards at grades C and D is important precisely because it is not, in the main, about poor marking. Rather, it is essentially about statistical modelling of ‘standard-setting’ (though the placement of

grade boundaries), a process that is completely removed from the local and personal distress regularly caused to large numbers of students, their parents and teachers by poor marking.

Yet it is the unglamorous, un-newsworthy business of addressing poor marking, year-in and year-out, that is the chief malaise of English examinations.

Local example

The case study school featured in the **Appendix** to this report is unusual only insofar as it has presented so simply and yet graphically the scale and range of the problems that poor marking generates for all schools in England.

The school is the size of an average academy school in the maintained sector. Between 2007 and 2012 the school felt impelled on 48 separate occasions to investigate further clear examples of poor marking and its effects on the grades awarded to specific groups of students. 18 of these cases related to GCSEs and 30 to awards at AS or A2. A spread of 19 separate subjects was involved.

When forwarding his school's data to HMC the comment of the head was pithy and to the point: *'Our return, for the record. A sorry picture emerges and this is the record of only one school!'*

This encapsulates both the vigilance of our schools and their determination to pursue justice for students, as well as scale of the challenge that confronts less well-resourced schools and colleges, including those in the maintained sector.

But more to the point, why in the first place should any school have to divert such energy to this a task?

Failing 6

ERRATIC AND INCONSISTENT MARKING AND RE-MARKING ON A LARGE SCALE IN RELATION TO A SPECIFIC EXAMINATION

What is the problem?

Each year it emerges in the weeks after results day that something is seriously wrong on a significant scale with marking in one or more particular examinations. Each year the identity of these defective examinations will vary by board, subject and level of study. Schools can be confident that such problems will arise but, initially, they cannot know that their local problem is one that will turn out to be much more widely shared.

Headline findings

- Putting GCSE English on one side (the problem there in 2012 is mainly related to grading across boards rather than marking within a board), the principal examination this year that so far has proved defective is the AS/A2 history A level provided by OCR.
- 21 HMC schools (approximately 1 in 5 of those entering candidates) have reported serious problems with this examination as at 24 September 2012. Up to this date, problems were being reported daily.
- Other schools report no problems with this examination. This compounds the confusion and dismay of teachers, candidates and (very often) universities.
- The range of problems identified for the examination is extremely wide and encompasses most marking, moderation, re-marking and re-moderating procedures.
- HMC has not yet explored thoroughly other examinations which may fall into this category in 2012. OCR history A level created an early lead this autumn, but there is the possibility that at least two other A levels may come to rival it over the next month in the scale of problems eventually reported.

Wider implications

Because it usually takes some weeks for the full picture about a problem examination to become clear, this is especially corrosive for students who need a particular grade in that subject to secure progression to post-16 study or to higher education.¹²

HMC is able to network intelligence among schools about each year's 'problem examinations'. Schools and colleges that do not have good networks of subject

¹² The situation in respect of university entry is especially messy when an exam proves as defective as did OCR A level history this summer. Some HMC schools reported that their students definitely lost their preferred place due to missing their required history grade, while others reported that their candidates all eventually secured their places, despite missing their history grade, because intelligence shared between *some* university admissions tutors in different institutions identified in time a wide-scale problem with this particular examination. This is as unjust as are the cases every year where students miss their required grade, and thus their preferred university place, due to being the isolated victim of a 'rogue marker' behind whom the relevant board subsequently stands at the appeal stage.

specialists in other institutions may not realise that their local problem is widely shared and needs to be pursued actively.

What specifically has gone wrong and what is the evidence?

As there is a choice of paired papers at AS and two mandatory units at A2, a total of six examinations is provided by the exam board. In 2012 HMC schools have reported problems with all six. Two of the six have a cluster of complaints. Meanwhile, some schools with concerns in one area report that the examination of other units has been unproblematic.

From this we can conclude that there is some very poor marking occurring on a significant scale across both AS and A2 units and that the effect of this – or not – nationally across those schools entering candidates for this examination is random and haphazard.

Where HMC schools have serious concerns about OCR AS/A2 history this year, the nature of these is very varied, falling under a range of broad headings that encompass most marking, moderation, re-marking and re-moderating procedures, as follows.

- Wide variation in the AS unit marks achieved by individual candidates whose teachers report that they were equally well-prepared for the two examinations.
- Concerns that initial marks were not only much lower than expected but clustered in the mid-range of the mark band for that grade (so making a revised grade on re-marking less likely).
- Schools used to ranking their candidates very accurately found that the grades awarded confounded these predictions very considerably.
- Close examination of marked scripts returned to schools revealed serious misapplication of the mark scheme.
- A2 coursework marks awarded by teachers who in previous years have been commended for accuracy of local assessment, reduced by the external moderator in a manner considered drastic and punitive.
- Suspicion that it is OCR's policy only to increase the mark of students whose original mark was assigned to the wrong grade band as opposed to the wrong mark (in some cases one or two marks are the difference between receiving a revised overall subject grade).
- Significant range in the re-marks awarded, most not shifting at all but some moving by up to 13 raw marks.
- In schools where the external moderation of the school's marking of its students' A2 coursework unit has been reported as '*having an excellent grasp of the mark scheme*', teachers discern that the re-marking of scripts at the board failed to adhere to the mark scheme.
- Dissatisfaction that full re-marking or re-moderation of papers actually takes place. In line with the Ofqual Code of Practice, this OCR service is formally described as a 'review of marking' rather than a re-marking exercise. The Board will not clarify upon enquiry from schools what processes actually occur and this obfuscation erodes professional confidence in the process on the part of teachers in schools.

Local example

To illustrate this range of concerns about OCR A level history in 2012, the following commentaries are taken from the reports of schools received by HMC over the last month.

Wide variation in the AS unit marks achieved by individual candidates whose teachers report that they were equally well-prepared for the two examinations.

'Last year this Centre achieved 28 A grades and no-one scored below a grade C. This year, with the same two teachers and similar cohort of students, we achieved 2 A grades, 3 E grades and 13 U grades.'

00000

'Results this August have been drastically out of line with expectations and with those of comparable students in previous years, despite the students having been taught by experienced and knowledgeable teachers with extensive track records of success in this unit. To an extent, some variations each year, of course, are to be expected, but, as explained below, these are far too drastic to be accounted for by normal annual fluctuations.'

00000

*'As explained, the teachers are experienced and knowledgeable, have achieved excellent results in the past, have attended several OCR briefing / training events and gone through every relevant Examiner's Report with a fine tooth comb. They are fully aware of the demands of this paper. The students were generally highly conscientious and with considerable ability (most had A*s or As at GCSE and were able enough to achieve on average one-and-a-half grades higher on another unit (F962B)). However, as a result of two consecutive years of drastically low results, students, parents and our management have little confidence in this unit and the way it is marked.'*

00000

'F961 is a concern. This was our strongest unit a year ago but has suddenly dropped to our weakest by at least a grade on average per pupil. There has been no change to the teacher or delivery. I fear it is being used as the unit to hold down grade inflation. If so, this is, of course, an arbitrary measure for doing so. Consistency is the key issue: just when you think you have a unit sorted, another goes haywire.'

Concerns that initial marks were not only much lower than expected but clustered in the mid-range of the mark band for that grade (so making a revised grade on re-marking less likely).

'The average obtained on F961 this year was 66.3/100. This is in striking contrast to the average for the F964 paper taken this year by the same boys, where the comparable figure is 82/100.'

The average of 66.3/100 is considerably out of line with the average obtained on this paper in previous years - the average mark obtained on F961 in 2010 was 78.2 and 84.5 in 2011.'

From the above it follows that there will have been differences between marks achieved by the same candidate on each of F961 and F964. In some cases these are extraordinary: 84 and 36; 80 and 41; and 88 and 27. It is also noteworthy that there are hardly any candidates close to the various grade boundaries’.

Schools used to ranking their candidates very accurately found that the grades awarded confounded these predictions very considerably.

‘The order seemed very surprising. One of our weaker candidates attained 108/120, whilst the most able were not being rewarded with the 90% + marks when they were expected to, in order to attain A grades. There is a sense from the grades given that the marking rewards style over substance and the most able candidates are therefore taking a risk by not moulding responses around the prescriptive mark schemes.’*

ooooo

‘We have an inaccurate rank order on F966 with clever, subtle and/or unusual answers awarded “b” or low “a” and plodding, jumping-through-hoops answers awarded a.’*

A2 coursework marks awarded by teachers who in previous years have been commended for accuracy of local assessment, reduced by the external moderator in a manner considered drastic and punitive.

‘Internal moderation in our school that has been considered “consistent with nationally agreed standards” for many years up until January 2012 is now deemed inaccurate by up to 20% when the course, internal moderators, mark scheme and assignment have not changed.

Notably in one case we have a student who completed his coursework for submission in summer 2011. This work was internally moderated as part of the 2011 cohort, but then this student became ill and opted to submit the work in summer 2012 instead. No change was made to the coursework, marking or moderation between summer 2011 and summer 2012. The 2011 marking and internal moderation was considered “consistent with nationally agreed standards” and would have awarded the coursework 62/80. In 2012, the assignment was moderated down by 17½% to 48/80.’

ooooo

‘This year a colleague and I marked the A2 coursework in the same manner as we have over the previous two years but to our surprise the marks were adjusted by the moderator by 25%. These included the scripts of [...] to whom we originally awarded full marks. In the end he was awarded a B grade for his coursework, which is hard to understand when it is remembered that he was awarded an A for his extended project and has been accepted to St. Hugh’s College, Oxford to study History.’*

Close examination of marked scripts returned to schools revealed serious misapplication of the mark scheme.

‘Students who attained A grades across the other units attained D and E grades in this paper. Having read the scripts, there seems no correlation between the grade awarded and the mark scheme, or any evidence of the examiner rewarding

the analysis of well selected knowledge. The key issue of well selected knowledge not being rewarded raises concerns about the knowledge of the examiner’.

Dissatisfaction that full re-marking or re-moderation of papers actually takes place. In line with the Ofqual Code of Practice, this OCR service is formally described as a ‘review of marking’ rather than a re-marking exercise. The Board will not clarify upon enquiry from schools what processes actually occur.

‘The Review of Moderation Report states on its front cover that it has two purposes, the first being “to justify marks awarded”. This should not be the purpose of the process. The Review should assess the OCR moderation in the light of concerns raised and establish whether the moderation has been accurate. The implication is that there was never any intention to question whether the mark scheme had been correctly interpreted by the moderator. The intention, it seems, has simply been to “justify marks awarded”.

00000

‘Over and above our concerns about securing accurate marks, we have serious concerns about the facelessness and lack of accountability of OCR. Who are the markers? Are they historians? Are they teachers? Do they teach the period they mark? Who undertakes the remarks?’

In short, OCR frequently gets it wrong the first time around and then is charged with investigating itself when criticized – at which point it usually admits either no fault or insufficient fault, thus meaning that the complaint cannot be prosecuted any further. It is an antagonistic and stressful circumstance, entirely corrosive to the confidence of those involved in teaching and learning.’

00000

‘My largest concern is the process of review/appeal. The re-marker gets the script with all the first marker’s comments and usually simply ticks them. The board refuses to EXPLAIN their marking. I had an issue two years ago with a student whose script was clearly grossly under-marked but was not changed in a remark. A letter from our school’s head asking for the marking to be explained was met by a refusal.’

Obstructions to redress: re-marks and appeals

Our final evidence relates to the highly unsatisfactory state of the re-marks and appeals processes that the exam boards are permitted follow.

Failing 7

SIGNIFICANT DOUBTS ABOUT THE ACCURACY AND FAIRNESS OF THE APPEALS PROCEDURE

What is the problem?

The formal procedure by which schools can challenge the marks or grades awarded to students is overseen formally by Ofqual in a very cursory manner.¹³ It is fundamentally flawed due to the way in which it allows the exam boards to hide behind protocol rather than requiring them to demonstrate, in an open manner, that scripts have been re-marked or re-moderated by an independent subject expert (rather than merely 'reviewed').

Headline findings

- Most HMC schools that go to formal appeal – either within exam board procedures or, as a last and expensive resort, to the Examinations Appeals Board – come away fundamentally dissatisfied with the process they have experienced.
- All comment that the boards are able to conduct the appeals processes they have devised in a highly defensive and opaque manner.
- Most also report exam boards 'putting down the shutters' (a phrase used more than once) so as to attempt to keep schools at bay and avoid incurring costs.
- Occasionally, when appeals are upheld there are far-reaching consequences. In 2009, the complaint of a single HMC school eventually exposed incompetence across of the AQA Philosophy A level examining team and led to the departure of many senior members of the examining team.
- Much more commonly, schools give up at various stages of the appeals process when it becomes clear that only procedures are being reviewed.

Wider implications

The lack of faith that schools have in the fairness or openness of the appeals process corrodes professional respect and confidence in the probity of the exam boards and their effective regulation.

¹³ See GCSE, GCE, *Principal Learning and Project Code of Practice*, Coventry: Ofqual, May 2011, pp. 50-53:

What specifically has gone wrong and what is the evidence?

Evidence has already been encountered (AQA GCSE drama (p.12 and OCR A level history (pp. 22) of the disillusionment of schools following an encounter with an appeals process that they come to see as thoroughly unsatisfactory and, in the words of one school, 'truly wretched'.

Most schools in England do not feel they can afford to take their case through all of the stages of appeal and, of the very small number that do, a large majority fall at the final hurdle in front of the Examinations Appeals Board (EAB).

David Dunn, is the head of Yarm School which was successful at an EAB hearing in 2007. He comments: *'Our then Director of Studies oversaw the appeal together with our then Head of Latin/Classics. They took the fight all the way and eventually won through sheer stubborn tenacity. Essentially, they knew they were right and so I allowed them time and resource to fight the battle in a way most wouldn't feel able to do. It became a matter of principle.'*

In contrast, St Albans School has lost two recent appeals, in 2008 and 2010. Headmaster Andrew Grant considers that the first case foundered on the lack of any independent subject expert able to back the school's claim that its candidates' work was qualitatively similar to, or superior to the exemplar material in the course textbook.

In the second case, St Albans' Director of Studies Mark Davies is of the view that *'HMC is absolutely right in saying that the appeals process allows exam boards to hide behind protocol. In our case our appeal was based largely on the examiners not applying their own mark schemes and assessment objectives fairly in the case of several of our candidates, but the EAB panel contained no subject specialists and eventually our case came down to a disagreement with the Chief Examiner with no independent person able to adjudicate between us and him. Even if a subject specialist such as a university academic had been called in, the exam board would have argued that that individual could not comment fairly on the marking because they were not at the standardisation meeting when the way in which the mark scheme and assessment objectives should operate were discussed and standards and expectations were set.'*

Local Example

At first sight it might be argued that the foregoing cases reflect merely the opinion of the school. In order to show that flaws in the appeals process identified by St Albans School are well grounded, the vivid account by Christopher King, Headmaster and Chief Executive, Leicester Grammar School, is a fitting note on which to end this report.

This detailed reconstruction of the complexities, cost and opaque nature of the appeals process echoes many of the central key themes set out in earlier sections of our report:

- a school identifying marks at odds with those of previous years, despite stability in teaching and student profile;
- initial impressions on the part of the school that papers had been poorly marked;

- an exam board that initially defended its ground on the basis that procedures had been followed, declaring ‘no further action is required’ even though internal papers revealed it considered that 40% of the candidates were worthy of a higher grade;
- insistence by the school that the original scripts should be seen to be re-marked;
- a Stage 2 Appeal Panel, none of whom were subject specialists, but who declared that the Ofqual Code of Practice had been followed leading to the case being dismissed, even though evidence had emerged of sub-standard work by the original marker and the team leader;
- a final appeal to the Examinations Appeal Board which rejected the school’s case because ‘the mark scheme was clear’, the board had fulfilled its own procedures and, in any case, the panel could not judge the quality of the original marking because it comprised non-specialists.

Appeal to the AQA Board by Leicester Grammar School re: GCSE English A3702/2H results, 2009

In 2009, my school appealed with regard to the AQA English grades of our 97 pupils who made up the Year 11 cohort. On results day it was clear to us that the pupils’ grades were, in almost all cases, lower than predicted and lower than they had achieved in the other subjects they had been examined in that year. The results were at odds with previous years’ grades and subsequently with the results gained by another similar ability year group in 2010. It is important to note that the ability of that year group in 2009 was very typical of any other Leicester Grammar School cohort which is a school where over two thirds of the pupils gain either an A* or A grade in their GCSEs. In addition, there had been no changes to the English Department in terms of ethos or personnel, all of whom were very experienced secondary school specialist teachers. We had been linked to the AQA Board and its English GCSE examination for about 15 years

Initially, we requested some selected answer papers be returned to the school where they had marks just below a grade boundary to minimise the possibility of a lower mark being awarded if a remark was sought. Our immediate impression was that the papers for those pupils had been poorly marked and a remark was sought. This brought about several improved grades and so we then entered the AQA Board’s appeal process.

We understood that Stage 1 of the Appeal process monitors whether procedures have been correctly followed. The conclusion reached by the Appeals Officer was that correct procedures had been followed. Frankly, this decision was no surprise to my school as our case had never been that the papers were not marked following Board procedures, but rather that they had been marked badly, inconsistently and inaccurately to the detriment of a significant number of our candidates.

The appeal review, attached to a covering letter from Dr Spencer at the AQA Board, referred to the original marker as being judged to be “sound and reliable although slightly lenient towards candidates in the 34-39 mark band”. The review of our candidates’ results was also carried out by an Assistant Principal Examiner whose work is judged as “consistently excellent”. The conclusion of the Appeals Officer was that “the correct standard of marking has been applied” and that “no further action is required”.

I confess this was the result I had anticipated but to my surprise the Appeals Officer had kindly supplied in tabular form the number of candidates who had their marks changed. In taking the Appeal Officer’s own statistics it revealed that, of the 15 candidates reviewed, 7 had their mark raised by 3 or more marks with one candidate being increased by 12+. It meant 46% of the candidates had seen an increase in their marks. Furthermore, the letter

reveals that 6 candidates had their grades raised, which represents 40% of the candidates reviewed.

Within the school, it was our view that this percentage of increased grades is about what we believed should have been the minimum proportion of higher grades which should have been awarded to our candidates. We did not understand how the Appeals Officer could possibly have been able to reach the stated conclusion quoted above, unless the parameters applied under Examination Board regulations were so loose as to be verging on the laughable.

Leicester Grammar School continued to follow the AQA Board's procedures, and despite our grave reservations about what would be achieved we appealed at the Stage 2 level.

The position we previously articulated to the AQA Board and the one which we certainly held to given the evidence above, was that all the entrants from Leicester Grammar School for this English paper should have been re-marked and reviewed. In common with other schools who find themselves in dispute with an Examination Board, we only asked initially for a re-mark to be carried out on candidates who appeared to be in a position where they could not go down a grade. It could well have been that other candidates deserved many more than twelve extra marks. It was our even more strongly held view that the marking had been inconsistent and that quite probably at least 40% of our whole cohort should have their grades raised. This re-mark should have happened without any risk to the candidates having their grades reduced from their present award level. In order to support this view we asked an independent academic from a university geographically unconnected with either Leicester or Leicester Grammar School to mark and review the work of the pupils. This process supported our view that the individual candidates had not received the correct marks and grades.

At Stage 2 we had learnt that the AQA Examination Board Appeal panel, none of whom were English specialists, although the Chief Examiner was present, would rely heavily in making their judgement on whether the Code of Practice had been followed. The Code of Practice was not made available to us until the School pressed for a copy to be made available. At this stage it was revealed that a remark had been carried out by an experienced assistant examiner, of all candidates who did not gain an A* grade. This review cast serious doubt on the marking of the original examiner, for which the Board then apologised. The team leader who had the role of reviewing the work of the examiner was also shown to be falling short of expected standards. The AQA Board then revealed that the principal examiner had also remarked and reviewed just 5 of the candidates work and pronounced herself content with the marks and therefore the grades awarded. The School pressed its case where it believed that the Code of Practice had not been followed but our case was dismissed.

Finally, we appealed to the Examination Appeal Board. It was stated by the Chair of the panel when we assembled for the final appeal hearing that the School should not understand that this was any court of law. It certainly felt as though this was the case as we had to provide 'evidence' before hand and we had to 'set out our case' and then be subjected to questioning. It certainly felt very much as though I, together with my Head of English, were having to defend our case not just state it. Our appeal to the EAB was rejected because 'the mark scheme was clear' and that the AQA had gone beyond required practice because reports had been prepared by the principal examiner on five candidates. They said no evidence had been provided of negative marking when in fact we provided a great deal of such evidence but the non-specialists had not an understanding of the subject to make such a judgement. The EAB recognised the original marking 'fell short of required standards' but they were sure the post-results procedures had been properly applied. At the EAB appeal meeting our examples of inconsistent and negative marking were ruled as inadmissible as the appeal panel judged that they were not qualified to judge the quality of such marking.

However, in rejecting our appeal the appeal panel went on to comment on areas which they said did not form part of their 'judgement'. They said the Board's Stage 2 report had not been adequately checked and several errors had been drawn to the attention of the EAB by the School. Secondly, that although the procedures for a Stage 2 appeal did not require reasons for the Board's decisions it believed it would be best practice to do so and would help centres.

The AQA was asked to give consideration to including justification for its decisions in future Stage 2 appeals. I do not know if this now happens.

In the final analysis, one of the Leicester Grammar School candidates improved the grade awarded from a D to a C, 17 moved from C to B, 7 grades from a C to either an A or A* and 16 from B to A or A*.

We were told this would have happened without any action by the school. We will never accept this and believe it was only because we fought through the protectionist barriers of the examination board, which was supported in the end by the toothless EAB, that any change took place. Not all our candidates got exactly what they deserved but we got closer than would have been the case without our efforts. Incidentally, the EAB appeal was heard in 2010 just before that year's Year 11 pupils sat their GCSE English paper.

Our case was that the English Paper was badly and incorrectly marked. We believe we proved that fact. We compared standardised scripts with our own candidates and found an A* grade standardised paper to have 31 mistakes but Leicester Grammar School pupil's papers with just 7 mistakes being awarded a lower grade. We lost our appeal at all stages because the examination Board said due procedure was followed. We did not say the papers were not marked, or the marks were not added up, but our case was that they were very badly marked and in the end the AQA had to apologise to us for getting it wrong. This was only after they had fallen into their default, defensive mode to protect the reputation of the Board. The review process is an enquiry of procedures and not of the essential quality of the fundamental marking of the pupils' work. At one stage part of the AQA case was that if they were doing such a bad job then why were so many more candidates being signed up for the GCSE examination. In stating this the truth was revealed that commercial priority overcame the moral imperative of ensuring that all candidates are awarded the grade their work deserves. We subsequently decided to move to the IGCSE examination for all our English Year 11 pupils with so far a highly satisfactory outcome.

C.P.M. KING
Headmaster and Chief Executive,

Appendix

Historical Evidence of Poor Examination Procedures At a typical HMC school

Inconsistent marking and grading in a range of subjects.

GCSE

| Year | SUBJECT | Unit | Awarding Body | Details of issue |
|------|--------------------|------------------|---------------|--|
| 2007 | Geography | 2403 | OCR | 8 out of 17 early EAR2 re-marks brought changes of up to 10 UMS. Extended review requested and granted. A few further students had their marks and grades advanced, one by +16 UMS. |
| 2007 | Drama | 3421/C 3421/W | AQA | Head of Department unhappy with results; mark scheme seemed “unfair”. No action taken, |
| 2009 | History | 1937 | OCR | 10 out of 13 early EAR 2 re-marks led to a subject grade change. Cohort re-mark granted. 1 further student: grade increase; 22 students: marks advanced. OCR: 2 lost scripts. Requested statistical increase for these candidates; agreed. One student’s grade change. |
| 2009 | English Literature | 2441 | OCR | 6 out of 12 early EARs showed large raw mark changes. Extended review granted. A further 21 students had their marks increased with two subject grade changes. |
| 2009 | English Literature | 2442 | OCR | One early EAR lead to a paper grade change from d to a*. Extended clerical check agreed by OCR; no further changes. OCR declined request for cohort re-mark; Headmaster appealed 6/11/09. No success. |
| 2010 | English | 2403 | OCR | Concerns over marking rejected by OCR; extended re-mark refused. |
| 2010 | History | 1937 | OCR | Poor marking. Cohort appeal granted by OCR. 22 students had their marks increased; two subject grade increases. |
| 2011 | English | 2431 | OCR | Evidence of poor marking, whole cohort re-mark for one paper. No appeals. |
| 2011 | English | 2432 | OCR | Cohort appeal requested 28/9 & agreed by OCR 7/10. 58 students had their marks raised (with many changes of subject grade). Decision to complain to OCR. |

| | | | | |
|------|--------------------|---------|---------|--|
| 2011 | French | 5FR04 | Edexcel | 18 out of 20 EAR re-marks led to an increase (by up to 15 UMS). Requested a full cohort re-mark. 34 students had their marks raised by up to 13 UMS |
| 2011 | Geography | 0460/12 | CIE | Department unhappy with initial marking. No action taken. |
| 2011 | Religious Studies | 40554 | AQA | Significant evidence of poor marking. Cohort re-mark was requested. Refused. AQA agreed to re-mark a further 10 candidates. Significant changes in marks from this re-mark. No appeal launched. |
| 2012 | Drama | | AQA | Head of Department unhappy with marking. Under consideration. |
| 2012 | History | 0470/04 | CIE | 53% A* in 2012; 63% A* in 2011. Very poor grades in CIE IGCSE History Paper 4 in 2012 - grade distribution suggests under-marking, but CIE's refusal/inability to provide individual component scores and copies of scripts and their slowness in responding to re-mark/review queries means that these disturbingly anomalous results are as yet unexplained. |
| 2012 | Religious Studies | 40554 | AQA | 4 out of 4 early EAR2 re-marks led to a change of subject grade; mark increases up to 14%. Director of studies recommends that an extension of the re-marking to the cohort be requested. |
| 2012 | Graphic Products | 45502 | AQA | Coursework marks were re-adjusted by between 25 and 30 marks per candidate. Cohort re-mark was requested. Department does not regard the significant marking adjustment as justified. |
| 2012 | Biology | 4BIO | Edexcel | 3 out of 7 early EAR2 re-marks led to a change of subject grade. Director of studies recommends that an extension of the re-marking to the cohort be requested. |
| 2012 | English Literature | A663H | OCR | Very inconsistent set of marks, reflective poor marking. Some unexpectedly remained stationary on re-mark, some raised by up to 15 UMS. |

A-Level

| Year | SUBJECT | Unit | Awarding Body | Details of issue |
|------|------------|-------|---------------|---|
| 2007 | Economics | 2882 | OCR | Unreliable marking; EAR2 re-marks led to significant changes in both directions. Requested an extended review of marking, which was rejected. |
| 2007 | Philosophy | PLY1 | AQA | A large upward change of 50 → 80 in Philosophy triggered AQA to initiate, at their expense, a cohort re-mark. |
| 2008 | History AS | 2584 | OCR | OCR lost script of a student. Headmaster appealed to OCR, rejecting OCR's calculated grade (using evidence provided by Head of Department). Stage 1 appeal granted. Mark raised: 78 - 86. |
| 2008 | History A2 | 2592 | OCR | Head of Department unhappy with published results. Headmaster wrote to OCR on 22/8 to apply for protected cohort re-mark; rejected by OCR. Requested Stage 1 Appeal on behalf of one candidate on the grounds of inappropriate application of mark scheme; further 5 candidates later added to this appeal. OCR agreed to have the Independent Essays of these 5 re-marked by Senior Examiner; mark changes: +11, stet, stet, +7 and +4. OCR extended this re-marking and another 8 candidates had their marks raised by up to 8. Stage 2 appeal (Headmaster and other staff). 14 more candidates had their marks raised by up to 10 (raw). |
| 2008 | English AS | 2707 | OCR | 5 out of 11 early EAR2 re-marks showed significant changes, in one case by 25 UMS. Extended review agreed. No further changes. |
| 2008 | Biology AS | 6103 | Edexcel | Head of Department unhappy with marking of Biology coursework. Extended review granted. Marks sustained. |
| 2009 | French AS | FRE2T | AQA | Unhappiness with initial marking. Head of Department requested return of oral tapes. EAR 2 re-marks led to some mark increases. In due course, AQA decided to increase the marks of <i>all</i> other candidates by 3; accepted by Head of Department. |
| 2009 | Spanish AS | SPA2T | AQA | Unhappiness with initial marking. Head of Department requested return of oral tapes. Alternative data provided; this satisfied Head of Department. No further action. |

| | | | | |
|------|------------------------|-------|---------|--|
| 2009 | Biology AS | 6BI03 | Edexcel | 4 out of 6 EAR 2 re-marks showed a change. Extended review initiated by Edexcel. 15 further students had their marks increased by up to 5 UMS. |
| 2009 | History A2 | 2589 | OCR | 5 out of 6 early EAR2 re-marks led to a change in paper or subject grade. Cohort re-mark granted. Marks of 6 candidates raised by extended re-mark. Requested scripts to service a possible further appeal. Arrived; Director of Studies advises no further action. Headmaster has given notice of appeal; Director of Studies advises this be withdrawn. Finished. |
| 2009 | History A2 | 2592 | OCR | 6 out of 8 early EAR2 re-marks led to a change in paper or subject grade. Cohort re-mark granted. Extended re-mark led to an increase in marks for 13 candidates. Headmaster wrote to OCR about pattern of mark changes. (13 candidates had UMS raised by 3 - 8; none had 1 or 2 mark changes). Scripts requested as evidence. Notice of partial Stage 1 Further Work appeal (one candidate) sent to OCR. Appeal rejected in Feb. 2012. |
| 2009 | Further Mathematics A2 | 7892 | OCR | Complaint by school: Marks of one candidate wrongly aggregated. Further maths grade raised from B to A on 22 nd October, by which time student had lost his University place. |
| 2010 | Economics AS | F581 | OCR | Initial marking unsatisfactory. Cohort re-mark granted. 3 students had their marks raised (by up to 12 UMS). The marks of a further 81 were sustained. |
| 2010 | Economics A2 | F585 | OCR | Concerns over marking rejected by OCR; extended re-mark refused. |
| 2010 | English Lit AS | F661 | OCR | Incredibly bad marking. Plethora of re-marks saw boys raised from D to A grade. Cohort re-mark granted; this led to increased marks for 5 students by up to 17 UMS. No appeals. |
| 2010 | Geography | F763 | OCR | Concerns over marking rejected by OCR; extended re-mark refused. |
| 2010 | Philosophy AS | PHIL1 | AQA | Very poor initial marking. Cohort re-mark granted. 13 students had their marks raised (by up to 16 UMS). |
| 2010 | Spanish AS | SPA2V | AQA | Department unhappy with initial marking. Cohort re-mark granted. 11 students had their marks raised (by up to 5 UMS). |
| 2011 | Economics AS | F581 | OCR | Evidence of poor marking, but insufficient to request a cohort re-mark. |

| | | | | |
|------|--------------------------|----------------|---------|---|
| 2011 | Economics AS | F582 | OCR | Early re-marks gave evidence of significant poor marking. In the time it took to request an appeal the remaining re-marks meant that the JCQ requirements were no longer met. 7 out of 21 significant changes. Cohort appeal requested 28/9 but rejected. |
| 2011 | Government & Politics A2 | F855 | OCR | The scripts of eight students were considered. No changes were made. Evidence of poor marking. |
| 2011 | History AS | F964 | OCR | Department unhappy with initial marking. Cohort re-mark granted. 5 students had their marks raised by up to 8 UMS. |
| 2011 | Spanish A2 | SPAN3 | AQA | Outcomes of several re-marks showed poor original marking. Insufficient evidence to permit a cohort re-mark. An attempt to request one was rejected by AQA. |
| 2011 | Spanish A2 | SPA4T | AQA | Outcome from Cohort appeal was that three students had their unit grade advanced. No appeal. |
| 2011 | Spanish AS | SPAN1 & SPAV2V | AQA | Outcomes of several re-marks showed poor original marking. Insufficient evidence to permit a cohort re-mark. An attempt to request one was rejected by AQA. |
| 2011 | Biology AS | 6BIO3 | Edexcel | In addition to the Service 2 re-marks, an additional 30 students had their marks raised by up to 9 UMS. There were two changes in subject grades. The outcome was still not acceptable but no appeal launched. |
| 2011 | Chemistry A2 | 6CH05 | Edexcel | Poor marking but no clear trend. |
| 2012 | Art | | Edexcel | The 2012 results show that the mark scheme was dishonoured; marks were lowered or the boundaries rose irrespective of the mark/descriptor. Moderators had clearly been told to support the shift they knew would occur by saying to lead teachers that AO3 had been poorly applied - even though the exam board descriptor clearly explains how AO3 is applied and we follow that rule. |
| 2012 | Mathematics AS | 4732 (S1) | OCR | This year the Statistics 1 Paper (OCR) was unusually difficult. When several members of the Department worked through it they all got wildly different answers. Although most of our pupils actually benefited we were alarmed to see the grade boundaries plummet to ridiculous levels, indicating that the board did not feel that the paper was fair. A pupil had to score 62/72 to get 100 UMS - a mark that would ordinarily get about 85 UMS. |

| | | | | |
|------|------------|-------|---------|--|
| 2012 | Italian AS | 6IN02 | Edexcel | Stage 1 appeals requested by Head of Department on behalf of two candidates. |
|------|------------|-------|---------|--|

Year-on-year variations in grades awarded

| SUBJECT | Awarding Body | Details of pattern / problem |
|-----------------------------|---------------|--|
| GCSE Drama | AQA | Irregular pattern over last four years. Significant concerns in regard to quality of marking. Inconsistency of grades being awarded. Outstanding pupils not recognised. |
| GCSE French | Edexcel | Irregular marking and awarding of grades. 2011 results incredibly low and pupils not awarded A*. Grade boundaries set very high. Moderation of internally assessed work significantly reduced. 2012 results over expectations. Pupils over awarded on marks and grades. Equally concerning as under marking. |
| GCSE English Lit & Language | OCR | Year on year inconsistent marking. Mark schemes applied very rigidly and firmly preventing good work being recognised. |
| GCSE Religious Studies | AQA | Year on year irregularity of marking. Remarks are being increased by large amounts and every year we are qualifying for a cohort remark. |
| A2 History | OCR | Consistently poor marking. Outstanding pupils are not recognised and marks are always not what we expect. The performance on different papers is always irregular which is not easily accepted. |
| AS Chemistry | Edexcel | Edexcel had to admit that the exams were too hard and too long (especially the multiple choice questions were too complex for a single mark. |

Significant doubts about the accuracy and fairness of the appeals procedure

| SUBJECT | Awarding Body | Details of pattern / problem |
|--------------|---------------|--|
| Biology A2 | Edexcel | 2012 - Paper was significantly too long and the majority of the boys failed to finish. Very able boys who should have gained an A* were prevented from this by not being able to complete the paper. Letter sent, acknowledged but rejected. |
| Biology AS | Edexcel | Coursework element of paper 6BIO3. The results of this were extremely bad in 2011, despite staff attending every training session offered. Some students took advantage of the service which informally assessed work prior to completion offered by Edexcel. These students gained 'Excellent' and then received a C grade on submission. 2012 Chief examiner was brought into School at great expense for a day. He trained every single member of staff in how to conduct, complete and mark this paper. Results were equally bad in 2012 despite this. |
| Chemistry A2 | Edexcel | One student requested a re-mark. On return of the re-mark he had been raised 9 marks (he required 10 to gain the grade which |

| | | |
|------------------------|-----|---|
| | | would give him access to university). On examination of his scripts there were three clear locations where the mark scheme had not been adhered to. A minimum of five marks had been omitted. Appeal launched in August. Return of appeal would not be until 21 days later. University place lost due to timescale. |
| IGCSE History | CIE | No disclosure of marks for individual components makes advising candidates on re-marks and appealing against re-mark outcomes with supporting evidence very difficult. |
| IGCSE Geography | CIE | No disclosure of individual unit marks. This makes internal analysis of performance impossible and cannot inform future teaching. This is exacerbated by not awarding the A* at unit level. This means a pupil can gain 81% and be awarded an A whilst a pupil can gain 100% and still gets an A. No way of centre being able to see variation. |

13th September 2012

