

GCSE English Awards 2012: A Regulatory Report



August 2012

Ofqual/12/5211

Contents

Foreword by the Chair and Chief Regulator	2
Executive summary	4
How we have approached this work	8
Standards in GCSE English in 2012.....	8
The awarding process for English GCSEs in 2012.....	11
Pressures on schools and colleges	17
Conclusions and next steps.....	20
Appendix 1: Background to the 2012 results	24
Appendix 2: Summary of GCSE results over time	27
Appendix 3: Maintaining standards – details of this approach.....	29
Appendix 4: Summary of meetings and actions over the last week.....	34
Appendix 5: Membership of the Ofqual Standards Advisory Group.....	37
Appendix 6: Glossary	38

Foreword by the Chair and Chief Regulator

We present our initial report on this year's GCSE English results in England.

We are grateful to everyone who has provided information and helped us, at very short notice. This list is long and includes many teachers, schools, colleges and their representative groups, particularly the Association of School and College Leaders (ASCL), as well as all the GCSE exam boards.

Ofqual was created in 2008, and became fully operational in 2010. We have the statutory responsibility for maintaining standards in qualifications, including standards over time. We bring rigour into the control of qualifications standards, using the strongest technical approaches in the field, agreed by experts. It has taken time for this rigour to feed through the system as it should: the implementation of these policies began with AS levels in 2009, with A levels in 2010, with most GCSEs last year and with English and maths GCSEs this year.

A level results stabilised in 2011 – there was little year-on-year change at national level – and they stayed steady this year. Those results have been accepted, and indeed welcomed by many. GCSE results stabilised this year, after two decades of grade inflation, and the reaction has been quite different. Many schools and colleges have been surprised by GCSE English results that have been well below their expectations. This led some to believe that the standard has been reset without their knowledge.

In this initial report readers will see that stable A level and GCSE results have been achieved through painstaking and expert work in exam boards over the last three years. However, the main focus of this report is the school concerns about GCSE English results this year. We explain as far as we can at this stage the patterns of GCSE English results, and why we think they have come as a surprise to some schools, colleges and students.

Those unfamiliar with the present school and exam systems may be surprised to see just how complicated GCSEs now are. Schools and colleges are increasingly sophisticated in their efforts to make sure that every student capable of achieving GCSE English grade C does so.

It is so important to students' future prospects that they achieve at least grade C English if they can. It is also extremely important to schools and colleges, because they care about their students' prospects, and also because student achievement at grade C English is so central to how schools, colleges and teachers are themselves judged.

Maintaining standards, so that GCSE outcomes stabilise, does of course have consequences for how schools, colleges and teachers are assessed, and we have

drawn this to the attention of both the Chief Inspector for Schools and the Secretary of State for Education.

Maintaining standards is rarely straightforward. The three English subject GCSEs (English, English literature and English language) are all completely new qualifications. The way they are structured and the way they overlap are complicated. Maintaining standards in the face of this level of change is very difficult indeed, as this initial report shows.

Signed:



Amanda Spielman

Chair



Glenys Stacey

Chief Regulator

Executive summary

The 2012 GCSE results for England, Wales and Northern Ireland were announced by exam boards on 23rd August. The results were broadly in line with the results in 2010 and 2011.

In total more than 5 million GCSEs were awarded this summer. Amongst these a new set of GCSEs has been introduced – English, English language and English literature.

On 23rd and 24th August some schools and colleges brought their concerns about GCSE English and English language to us. We immediately started to look carefully at how exam boards managed the awarding of these GCSEs this year.

We are not aware of any significant issues for other GCSE subjects. Our review, reported here covers only English and English language.

The standard in English GCSE was maintained although changes in the mix of students meant that results were down overall by 1.5 percentage points. This was in line with expectations. That is not the issue. The issue is that for some schools and colleges their results are not what they expected and the differences in some cases were hard to understand.

These qualifications are different from previous English qualifications in a number of ways. Part of each qualification is done as a controlled assessment (coursework carried out under the supervision of a teacher), worth 60 per cent of the marks. The qualifications can be taken in stages or modules during the year, but at least 40 per cent of the assessment must be taken at the end of the course.

We secure the standards of qualifications. If we were to compromise on this, we would undermine the integrity and rigour of the qualifications we regulate.

In maintaining qualification standards across exam boards and over time we consider fairness to students past, present and future.

We set out in this initial report what we think people need to know about qualification awarding in general, and how it worked for GCSE English subjects this year. We explain what we have been doing to understand and evaluate the concerns being expressed about the GCSE English subjects results, outline what we have found, and what we intend to do next.

The GCSE English suite was first taught in 2010, and awarded in 2012, with the inevitable challenges any new qualifications bring to teachers, schools and examiners. Examiners, of course, are used to dealing with new qualifications.

For GCSE English this summer a complex set of circumstances came together to create an unusual situation for schools, colleges and their students.

On the basis of the work we have done so far, our initial findings are:

- **The standard set for these English GCSEs is comparable with the standard in previous years.** That means that a candidate awarded a C grade this year achieved the same standard as a candidate awarded a C grade last year in the same subject.
- **On thorough investigation we can confirm that each exam board set standards for the qualifications in the manner expected and at the appropriate standard.**
- We have looked carefully in each exam board at how their examiners set the June 2012 grade boundaries for all units. We found that they acted properly, and set the boundaries using their best professional judgement, and taking into account **all** the evidence that was by then available to them. We have seen evidence of how it was done. **The June boundaries have been properly set, and candidates' work has been properly graded.**
- **The issue is not June, but January grade boundaries.** A minority of candidates sat their units in January, and again we have looked at how grade boundaries were set then. Again, examiners used their best professional judgement, but they had less information to go on, less hard data to help them come to a judgement. This was both because most candidates were not sitting at that time and, because these were new qualifications, examiners could not rely absolutely on how standards were set in past years.
- **Three things came together, and made the job of setting standards difficult in January 2012:**
 - First, changes to the syllabuses themselves
 - Second, the nature of English as a subject. English subject examiners have found that setting standards in English, in new qualifications, is difficult.
 - Third, the structure of the qualifications. They are made up of a combination of modules **and** they have a high proportion of controlled assessment (60 per cent).

In short, grades awarded for the June modules were right, but it is hard to square them with the January results.

- There were different written papers set in January and June for the AQA foundation tier. It is wholly appropriate that the grade boundaries for these papers were different. For some of the controlled assessment units, principally Edexcel and AQA, the grade boundaries did move between January and June for the same pieces of work. In retrospect the January grade boundaries were generous. This will have had a minimal direct impact because most schools submitted written controlled assessment in June. For example only 2 per cent of AQA candidates were awarded grades on their written controlled assessment in January. Most students took the written exams in June and completed their controlled assessment at the same time.
- **Some schools were over reliant on the January 2012 grade boundaries particularly in relation to written controlled assessment.** Exam boards published the grade boundaries set in January 2012, and although these had a caveat, many schools used these to set expectations. For schools, as for exam boards, this will have been their first experience of dealing with modularised GCSEs for most English departments. Both could have shared a better understanding of the use made of grade boundaries.

Looking forward

First we want to go through our analysis and evidence with the representative groups for schools and colleges, so they can see it for themselves. We also want to talk with schools, exam boards and assessment experts about the lessons to be learnt here and what can be done better in the future. One clear lesson is that if changes are made to what teachers teach and how it is tested, the maintenance of standards over time gets very difficult.

The GCSE system is complex. A system of this complexity is difficult for schools and colleges to understand, and it increases the risk of problems such as those we have seen this year. Modularisation creates a particular difficulty with maintaining standards in graded qualifications. We have already announced that we will be moving away from a modular system for GCSEs in England after the forthcoming school year. We believe that this year's experience shows that this was the right decision for students in England.

Students starting new English GCSE courses this autumn will not be doing modular exams. They will be preparing for exams and controlled assessment at the end of their course in June 2014.

Students who are part way through GCSE English courses should be reassured that when exam boards set grade boundaries in future they will be able to do so based on

a more complete understanding of how the qualification works and data from this summer. We will be monitoring the delivery of this qualification carefully. It is important that schools know what conclusions it is safe to draw – and not draw – from the information they receive from exam boards and elsewhere. Exam boards will review the advice and guidance given to schools about GCSE English, including; its structure, how grade boundaries are set and how grade boundaries may be used with caution in predicting student outcomes.

As for all GCSEs and A levels, schools can make enquires about results (known as EARs) to have the marking of an individual script checked. No doubt some schools will want to do this and many have already asked their exam boards to do this. EARs will not result in changes to grade boundaries. So some schools may wish to discuss with their exam boards whether their EARs are appropriate and whether there is further information the exam board can provide about performance at the school level to guide teachers. We know exam boards are ready to do this.

Any student disappointed with the outcome of his or her GCSEs has an option to do a resit. The next opportunity for this is in January 2013. Given the concerns of some schools and colleges about their students' qualification outcomes from June 2012, each of the exam boards offering GCSE English and English Language qualifications will provide an exceptional, one-off, re-sit opportunity in November 2012. The January 2013 and June 2013 assessment opportunities for these specifications will go ahead as planned. To do this we will need to relax some of the rules around resits. We would welcome views on the proposal by 7th September. Exam boards will confirm arrangements in the week beginning 10th September.

We set out in this report our decisions. These decisions apply to England only. There are separate issues in Wales and Northern Ireland which their governments and the regulators are considering, and we will discuss with them the conclusions that we have reached and the work they are doing.

How we have approached this work

Our aim in this work has been to:

- confirm and explain what has happened to GCSE English results this summer
- take swift action if problems are found
- maintain confidence in the exam system
- identify broader questions for further work by us and others.

To do this, we have gathered data and evidence from a range of sources to help us understand what happened with awarding. We have had meetings and input from exam boards, teacher associations, individual schools and others. More details of what we have done and the questions we asked exam boards are in Appendix 4.

In considering how to respond to the concerns of teachers and others, we have applied the principle that underpins all that we do. We secure the standards of qualifications. If we were to compromise on this principle, we would undermine the integrity and rigour of the system.

This means that we maintain standards: students should get the grade that their work deserves. When issues like this occur we need to maintain standards over time. We need to consider fairness to students past, present and future.

Standards in GCSE English in 2012

A new set of English GCSEs, first taught in September 2010, was awarded for the first time in 2012. There are three subjects, English, English language and English literature, together known as the “English suite”. Students in England chose either English, or English language and English literature.

Across English and English language GCSEs, the proportion of candidates getting grades A*–C fell from 65.4 per cent in 2011 to 63.9 per cent this year. This may suggest that the qualifications were made harder. However, our aim in setting standards year-on-year is to keep standards the same. The evidence suggests that we succeeded. This section explains this.

The principle applied in setting standards for a new GCSE is that a student should get the same grade as they would have done if they had taken the old version. We call this approach “comparable outcomes”. It aims to prevent “grade inflation” – that is, increases in the proportion of students achieving higher grades without evidence

of real improvements in performance. It also adjusts for the dip in performance that can arise when a new qualification is first taken. It is more sophisticated than “norm referencing” where fixed proportions of candidates would be awarded a grade without reference to candidate achievement. More detail of our approach to setting standards is in Appendix 3.

New modular GCSEs were awarded for the first time in most subjects in 2011, and we used comparable outcomes to help maintain consistency between years. The outcomes for the new subjects, where we were aiming for comparable outcomes, were broadly consistent between 2010 and 2011 – grade A* fell by 0.2 percentage points compared with 2010, grades A*–A fell by 0.4 percentage points and grades A*–C fell by 0.3 percentage points. Meanwhile in the other subjects, including English, grade A* rose by 0.4 percentage points compared with 2010, grades A*–A rose by 1.4 percentage points and grades A*–C rose by 1.6 percentage points.

The comparable outcomes process will produce similar results year-on-year if the cohort for the subject is similar, in terms of ability, in each year. The initial expectation for summer 2012 was therefore that the overall change in English results should be close to zero. We checked this by comparing the predictions made by the exam boards of results for each GCSE, based on prior attainment data, with the actual profile reported by the exam board.

English/English language – 2012 comparison of predicted and actual grade distributions

	AQA predicted	AQA actual	All boards predicted	All boards actual
A*	7,571	8,319	12,954	13,819
A*–A	39,528	38,690	66,167	65,778
A*–C	171,435	169,547	290,110	288,788
A*–F	249,687	248,698	421,984	421,018

Source: information from exam boards

At national level, predictions fitted fairly closely to actual results.

The headline grade A*–C figure fell this year by 1.5 percentage points from 65.4 per cent to 63.9 per cent. This movement related to the changes between 2011 and 2012 in the profile of candidates, rather than to a change in the grade C standard. Overall, as the table below shows, there were more candidates taking English and English language in 2012 than in the previous year.

English/English language 2011 and 2012 – numbers of UK candidates

	2012	2011
	'000	'000
English	171	650
English language	498	-
Total	670	650

Source JCQ

In the exam control processes, results are compared with prior attainment for “matched” candidates, that is, those for whom there are known Key Stage 2 results. These include most pupils in the state system. It is the change in the Key Stage 2 point score of these matched candidates that needs explanation. There were two significant changes to the mix of the English GCSE entry in 2012.

First, there were about 23,000 fewer candidates from selective and independent schools, about 3.4 per cent of the total, who will probably have migrated to the international GCSE or other qualifications. These candidates will typically have relatively high Key Stage 2 point scores, so their departure will have lowered the attainment profile of the cohort.

Second, there were about 43,000 more candidates from non-selective state schools and colleges, even though the number of 16-year-olds in the national cohort fell slightly. These extra candidates are 6.4 per cent of the GCSE entry. Some of these will be re-sit candidates, who generally have lower Key Stage 2 point scores than average. Others will be candidates who in the past would have been entered early for a winter exam: these early entrants also tend to have lower Key Stage 2 point scores than summer entrants.

So to summarise, for candidates with a given level of prior attainment at Key Stage 2, the comparable outcomes policy has ensured that their expected outcome has not altered. But because of cohort changes, the reported national result is down.

The awarding process for English GCSEs in 2012

Having established that standards have not altered at national level, we have reviewed exam board and regulatory processes for GCSEs in the English suite in January and June 2012 to identify any abnormalities in the details of the awarding. We address first, the June awarding and secondly, the January awarding, since some concerns relate to comparisons between January and June grade boundary marks.

We have reviewed in detail the awarding for the GCSE English suite by the four main exam boards (AQA, Edexcel, OCR and WJEC), but since public attention has focused on two AQA assessment units, the foundation tier written paper and the written controlled assessment, we set out details of our review of the awarding only of those two units, in June and in January.

Written controlled assessment submissions were overwhelmingly made in the summer. Across all boards 719,000 submissions were made in the summer, against 51,000 in January, which is only about 7 per cent of total entries. Only 2 per cent of AQA submissions were made in January.

AQA foundation tier written paper assessments were less skewed towards the summer: about 135,000 in January and about 298,000 in June, so just over 30 per cent of foundation tier entries were in January.

When exam boards set boundaries and award grades, they use qualitative and quantitative evidence. This includes samples of candidates' work, reports from senior examiners, grade descriptions and also statistical information about the groups of candidates ("cohorts"). They must also have reference to the comparable outcomes approach which aims to ensure that standards are maintained at qualification level (see Appendix 3). Despite the best efforts of senior examiners, written papers can vary in demand and boundaries are set each year based on a combination of senior examiner judgement and statistical data about how well candidates perform.

For summer awards, we monitor the awarding process, including attending some awarding meetings and reviewing data on outcomes, and this year we observed some unit level awards as well. When making unit awards during the early stages of a qualification, examiners' judgements have to be relied on more heavily, because much less evidence is available about the cohort's overall performance than when the whole qualification is being awarded.

Summer awarding

Based on our initial review, summer awarding was carried out properly and examiner judgements were made reasonably in all exam boards. The evidence from our

regulatory monitoring at the time and from our reviews over the last week show that the outcomes were appropriate. On that basis, we have no reason to ask the exam boards to revisit their awarding decisions.

While we discuss the AQA units in detail below, as they have attracted most attention, we did not find anything substantively different here from our review of any other exam board.

Examiners in all boards had to take account of a number of things in setting boundaries.

First, the awards that had already been made in completed units (modules). Students completed units at a number of points in the course. Controlled assessments could be taken in January 2011, June 2011 or January 2012, and awards were made shortly after they were marked. Students could, though, choose to submit a replacement assessment with the aim of getting a better result. Written exams are also available in January and June of each year.

Second, they looked at the units (both exams and controlled assessments) taken in the summer. Forty per cent of the assessment must be taken at the end of the course.

Third, they considered the expected pattern of outcomes at qualification level. The comparable outcomes approach requires exam boards to make explicit expectations for what the pattern of results will be, at qualification level. Balancing these factors is always a challenge when awarding modular, graded qualifications.

In reviewing the awarding process, we looked at both the process – how decisions were made – and the outcomes, the level at which grade standards were set and the awards that were made. We reviewed in detail the steps AQA took to set an appropriate grade boundary for each of the units, including looking at: the chair of examiner's reports and proposals to change grade boundaries, AQA's technical reports on awarding and the principal moderator's reports. We had already observed some of the stages of the awarding process first hand as part of our normal monitoring process.

It is apparent from the records of AQA meetings and decisions taken that examiners were aware of and tried to address the range of challenges they faced. Their difficulties were compounded by uncertainties about the relative ability of the cohorts for English and English language and the value of the predictor information that was available to them. They reviewed and recommended revising grade boundaries as more candidates' work became available to them during each awarding exercise.

The reports we looked at record that AQA staff had been at pains to explain to teachers that grade boundaries could change, but that teachers continued to press

staff for information about grade boundaries. Concerns are expressed in the records about teachers over-rewarding controlled assessments they had marked. This was explained in part by teachers' lack of familiarity with controlled assessment, which had replaced coursework. But the reports also record concerns that teachers' and examiners' expectations of speaking and listening quality, in particular, were sometimes misaligned.

We have looked unit by unit at AQA's approach that resulted in AQA's responsible officer signing off the awards for the January units on 22nd February 2012 (following an awarding meeting on 21st and 22nd February) and for June units and the whole qualification on 9th August 2012 (following an awarding meeting on 21st and 22nd July). The awarding meetings cover all the relevant units, which allows foundation and higher tier papers to be considered together.

We set out below the facts and history of awarding of each of the relevant units and discuss the main issues arising from each.

Foundation tier paper

The foundation tier written paper (unit ENG1F) was first offered in January 2011 and was offered in June 2011 and January and June 2012. In January 2012, 54,000 candidates sat the paper. In June 2012, 141,000 candidates sat this unit.

As is usual practice, for both series the mark schemes were finalised by the principal examiner for this paper at standardisation. For each series the principal examiner prepared a report for the awarding meeting on the way candidates had responded to the paper. This report for the June paper noted that: "the overall demands of the paper were very similar to previous series".

At the awarding meeting in July, senior examiners reviewed scripts at a range of marks and the chair's report notes that: "The tier F agreed C boundary, initially 52, was moved up to 53 on revisiting the [evidence] in the light of further statistical information. Although this mark was significantly higher than the mark for the reference year [June 2011] it was felt that this mark was a truer reflection of the quality of candidates' work in relation to the C grade criteria".

Controlled assessment unit

The controlled assessment writing unit (unit ENL03) was first offered in January 2011 and then in January and June 2012. In June 2012, 287,000 candidates entered this unit. The principal moderator's report for summer 2012 noted that the majority of folders [work] for controlled assessment were within the marking tolerance but had been over-marked by teachers, particularly at the grade C/D borderline.

The July awarding meeting considered a range of evidence. The initial recommendation for the grade C boundary was 47. However, in light of the evidence

of the level of performance seen in the archive work (from previous series), the chief examiner recommended a revised boundary mark of 46.

Following the awarding meeting, the outcomes were reviewed within AQA by senior technical staff. The approver's report noted that changes to the controlled assessment boundaries since the January series were likely to cause disquiet in schools, but also noted that the committee had been unanimous in endorsing the controlled assessment boundaries as there was "evidence of significant teacher over-marking". The boundary mark of 46 was confirmed.

Issues arising for June awarding

In the case of the AQA foundation tier paper, it was suggested that the change in the boundary mark between January and June was abnormal, and that the grading of the June candidates might have been over-severe to compensate for more generous grading in January.

In the case of the controlled assessment, there was clearly a risk of confusion from the fact that the June awarding meeting reached a different decision on the grade C boundary mark from the January meeting. The grade C boundary mark was in fact raised between January and June by all three of the boards (AQA, Edexcel and OCR) which had offered a January controlled assessment submission opportunity.

Out of 19 units offered by these three boards, in only two was the grade C boundary not higher in June than in January.

January awarding

We reviewed January awarding across all boards and English/English language units to ensure that all awarding processes were properly carried out and awarding judgements were properly made.

On controlled assessment, we found in all boards that grade boundaries were set properly in the context of small entries with limited historic and other data to provide context: it was only with hindsight that it was possible to see that these boundaries were too generous.

AQA foundation tier written paper

In January 2012, 54,000 candidates sat the foundation tier written paper. In June 2012, 141,000 candidates sat this unit.

The awarding meeting was chaired by the AQA Chair of Examiners for GCSE English.

The awarding meeting considered a range of evidence, including:

- reports from senior examiners
- the principal moderator's report which included information on how well teachers had followed the mark scheme
- appropriate samples of candidates' work; this is needed to give a sound basis to key grade boundary decisions
- detailed technical and statistical information, including the mean mark for the unit, the distribution of marks, entry types and information about how well candidates had responded to individual questions.

During the meeting, issues such as the contrasts in the entry profile and the mark distribution between this unit and the equivalent unit for GCSE English were explored.

The proportion of students achieving a grade C on the foundation tier paper dropped from 37 per cent in January 2012 to 10 per cent in June 2012. If the June results were right, that might suggest that AQA's award in January 2012 was too generous. However, the evidence that AQA had at the time – the data on expected awards – suggested that they might have been too severe, and the grade boundary that was set was higher than the data would have suggested. It is now clear that they had limited evidence from a relatively small entry, but our view is that the judgements made at the time were sound.

Following the January award, AQA's internal accountable officer's report noted that one of the emerging issues was the relatively low boundary marks (as a proportion of the maximum mark) and the fact that the grade C boundary marks were similar on both tiers. Since C is the highest available grade on the foundation tier, we would usually expect the C boundary on the foundation tier to be higher than the C boundary on the higher tier paper. The June paper boundary mark of 53 is much closer to what we would expect.

AQA controlled assessment – writing

In January 2012, 7,000 candidates entered this unit. A summary of the evidence considered in the awarding meeting is set out above. The boundaries were considered in the order of grade C, then A, and then F. For each of the A, C and F grade boundaries, controlled assessment work was reviewed by the senior examining team. The marks were out of 80. They reviewed work marked in the range 41 to 45 for grade C, 62 to 66 for grade A, and 14 to 18 for grade F.

Having reviewed the full range of evidence, for each boundary awarders came to a collective view on the boundary mark to recommend. At C, there was no change to the boundary mark (43) from the June 2011 unit, though there was a range of views expressed about this. The report from the AQA support officer noted that the marks for the controlled assessment appeared to be very high but that there were very low numbers of entries, which might cause difficulties when setting grade standards for the qualification as whole.

Our observation of the meeting concluded that it had been conducted in line with AQA processes and the relevant regulations – the recommendations were made with the aim of ensuring the comparable outcomes required at overall subject level.

Conclusions

Our review shows no grounds for re-opening the awarding, at AQA or elsewhere, of the foundation tier paper. The records of the awarding process show that due weight was given to the various factors we would have expected. In retrospect, the January grade boundaries were generous, but they were less so than the data had suggested they should be.

For the controlled assessment, the outcomes of the January awards were generous. Some of the small number of January candidates will have been over-graded on this module. With more evidence available in June, it was clear that the grade boundaries had been set too low, though in the time available, we have not been able to calculate the numbers of students affected. While logically these grades would now be adjusted downwards (and indeed this option was considered when June awards were being finalised), we have considered whether that would be the fairest option, and have concluded that it would be unfair to those candidates to take away the grade they thought they had earned, more than six months later. There is also no precedent in the exam system for such a retrospective adjustment to grades.

It is regrettable that the publication of grade boundaries for the January assessments could have led schools to assume that the boundary would remain constant, and we will review with the exam boards any lessons from this. We are working with exam boards to provide an additional early controlled assessment submission opportunity for candidates to retake this unit.

Our final two observations on the awarding process are these. First, the differences between the grade boundaries in different series arise from a system that is complex and difficult to understand. Maintaining standards will be much easier when we do not have modular GCSEs in England.

Second, had the grade boundaries for January carried through to June, there would have been very significant grade inflation at qualification level: there would have been a big increase in the proportion of candidates getting grades A*–C at English

GCSE which was not justified by the evidence. As the regulator, we could not have defended such an increase. It would have put us in breach of our statutory standards objective.

Pressures on schools and colleges

The pressures on schools and colleges have a significant bearing on school reactions to the GCSE English results. We have listened to schools and their representatives over the last week. We want to understand their concerns and perspectives, and to consider whether there are things we can do to respond to them. In this section, we discuss some of the issues that are faced by schools, particularly around predicting results.

All state-funded schools in England have a range of accountability measures reported in school performance tables. They are also expected to show improvement over time in the context of Ofsted inspections. Both of these are high stakes for the relevant schools.

Schools also are responsible for preparing students for the next stage of education. Admission to A level courses will often require a B grade or even an A grade at GCSE. Colleges may have a general admission requirement of five C grades. However, it is clear that the C/D borderline is the boundary of strongest interest in most schools, and that most schools therefore invest considerable resources in getting borderline candidates up to or just above this level. For schools whose intake is skewed towards lower prior attainment, a substantial minority of students are likely to be “borderline” candidates.

In practice this means that most schools have a very strong need to know “what grade C looks like” in the context of each type of assessment. To quote one school Head of English: “Central to the design of any new syllabus must be absolute clarity on the board's part over what standards of work look like. For example, what a C grade controlled assessment is or what a C grade performance in an exam looks like.”

Schools also use the pathways through qualifications that give the greatest possibility of achieving an overall C grade. In some cases candidates have completed controlled assessments in January, but not “cashed in” the grades, choosing instead to submit further controlled assessments in June, even where the grades have been as good as they could reasonably expect. One reason for this is to make it possible for candidates to re-sit the written paper in June in the hope of a better grade, but to keep the option of cashing in the earlier grade on the written paper while still being certain of fulfilling the requirement that 40 per cent of marks must come from a terminal assessment. The multiplicity of pathways means that candidate outcomes

have sometimes come to depend too much on school choices as well as their own performance. This is another aspect of the complex system that is in place.

As a new GCSE beds in, the collective understanding of C grade performance improves, and schools become increasingly confident about it. But in the early years, schools rely heavily on exam board sample materials and on the boundary marks in early sittings. Unfortunately this is the period in which boundary marks are most unstable as examiners find the right level at which to pitch and grade assessments. At this stage exam board sample materials are not real candidate work for the new specification: they are either written by examiners or borrowed from previous versions of the GCSE.

In this first year of awarding the new English specifications, schools have reported being heavily reliant on the grade boundaries from earlier units, probably more so than would normally be the case. For example, some schools made sure that students worked on controlled assessment preparation until they reached a mark that the school believed would correspond to a grade C.

Many of the discussions we have had with schools have centred on student and school outcomes relative to their predictions. Most schools track predicted outcomes: some from as early as the first externally assessed or moderated unit, some just for the last three months of the GCSE course. Schools submit predictions to exam boards in May, shortly before the final assessment.

Predictions may be made holistically, or they may be calculated very directly from student results on practice assessments in the context of exam board sample materials and previous grade boundaries. Predictions inform senior management of expected overall outcomes for their school. They are also used by teachers to decide whether a student should continue trying to improve their likely controlled assessment mark, and whether a unit grade should be cashed in or not.

The significant changes to the boundary marks in the two AQA units between January and June might therefore have contributed more than they normally would have done to creating a big gap between school predictions and their outcomes.

However this needs to be looked at in context. The ASCL has recently surveyed member schools on this year's AQA English outcomes. Schools believe that they are able to predict grades fairly accurately at cohort level if not at individual student level, and their representatives are suggesting that the gap between school predictions and this year's outcomes must indicate an exam board problem.

The following table records the disparities reported to ASCL in the past few days by just under 750 schools.

School outcome	AQA %	Total %
more than 10% worse than expected	42%	42%
5.1 to 10% worse than expected	32%	32%
1 to 5% worse than expected	14%	13%
Broadly as expected	8%	9%
1 to 5% better than expected	3%	3%
5.1 to 10% better than expected	1%	1%
more than 10% better than expected	0%	0%

Source: ASCL and OCR

Two things emerge from this data.

First, given that national results were only marginally down, it shows that schools in aggregate must have been expecting substantial rises in English GCSE outcomes this year. We know from exam board data that schools do tend to over-predict, but the school reaction this year suggests that the gap between expectations and outcomes was wider than usual.

We found some confirmatory anecdotal indications that school expectations of improvement in their English/English language grades were particularly high this year, perhaps in part because of the change in structure of the qualifications. As one senior English adviser put it, “we would expect students to do better than on the old specification if their controlled assessment was of sufficient quality”.

Secondly, in the ASCL survey, despite the media focus on AQA, the gaps between predictions and actual outcomes are very similar for AQA and for other exam boards..

We acknowledge that this year’s English results have come as a shock to some schools, and some of the school-level outcomes are hard to explain. We have received very little school-level data this week so more will be needed to resolve some of these questions. However in this report we have tried to set out for schools what happened and why, and to explain the decisions that we and the exam boards have made. We have also considered carefully whether there are any specific actions we should take in relation to this year’s candidates.

Conclusions and next steps

This work we have done over the last week has been a valuable exercise for us and the exam boards. It has forced us to confront some serious challenges to the way we have approached the maintenance of standards. It has shown how complex the system has become. It has shone a light on the way that schools engage with the exam system and the pressures they face. It has demonstrated that many people do not understand the system. Through this document we hope that those who are working with qualifications will be able to get a better understanding of how the system works and the reasons for the issues this year.

Overall, we think that our approach to qualifications has stood up well. Some of our hypotheses about what might have happened have turned out not to be the case. Exam boards – who have co-operated well with this work – have been able to demonstrate that they followed the awarding processes we expected, and that the standards judgements they made were robust.

We know that schools and colleges will be disappointed with our conclusions. Many will have been hoping that we would conclude that the June grade boundaries were too harsh, and that we should return to the January grade boundaries. We have considered this option, but as a standards regulator it would be impossible to justify doing so: it would undermine standards and damage confidence in the system. Nonetheless, we hope that schools and colleges will be reassured by the thoroughness with which we have approached this work, and the detailed explanations we have set out of what happened and why.

On the basis of the work we have done so far, our initial findings are:

- **The standard set for these English GCSEs is comparable with the standard in previous years.** That means that a candidate awarded a C grade this year achieved the same standard as a candidate awarded a C grade last year in the same subject.
- **On thorough investigation we can confirm that each exam board set standards for the qualifications in the manner expected and at the appropriate standard.**
- We have looked carefully in each exam board at how their examiners set the June 2012 grade boundaries for all units. We found that they acted properly, and set the boundaries using their best professional judgement, and taking into account **all** the evidence that was by then available to them. We have seen evidence of how it was done. **The June boundaries have been properly set, and candidates' work has been properly graded.**

- **The issue is not June, but January grade boundaries.** A minority of candidates sat their units in January, and again we have looked at how grade boundaries were set then. Again, examiners used their best professional judgement, but they had less information to go on, less hard data to help them come to a judgement. This was both because most candidates were not sitting at that time and, because these were new qualifications, examiners could not rely absolutely on how standards were set in past years.
- **Three things came together, and made the job of setting standards difficult in January 2012:**
 - First, changes to the syllabuses themselves
 - Second, the nature of English as a subject. English subject examiners have found that setting standards in English, in new qualifications, is difficult.
 - Third, the structure of the qualifications. They are made up of a combination of modules **and** they have a high proportion of controlled assessment (60 per cent).

In short, grades awarded for the June modules were right, but it is hard to square them with the January results.

- There were different written papers set in January and June for the AQA foundation tier. It is wholly appropriate that the grade boundaries for these papers were different. For some of the controlled assessment units, principally Edexcel and AQA, the grade boundaries did move between January and June for the same pieces of work. In retrospect the January grade boundaries were generous. This will have had a minimal direct impact because most schools submitted written controlled assessment in June. For example only 2 per cent of AQA candidates were awarded grades on their written controlled assessment in January. Most students took the written exams in June and completed their controlled assessment at the same time.
- **Some schools were over reliant on the January 2012 grade boundaries particularly in relation to written controlled assessment.** Exam boards published the grade boundaries set in January 2012, and although these had a caveat, many schools used these to set expectations. For schools, as for exam boards, this will have been their first experience of dealing with modularised GCSEs for most English departments. Both could have shared a better understanding of the use made of grade boundaries.

Looking forward

First we want to go through our analysis and evidence with the representative groups for schools and colleges, so they can see it for themselves. We also want to talk with schools, exam boards and assessment experts about the lessons to be learnt here and what can be done better in the future. One clear lesson is that if changes are made to what teachers teach and how it is tested, the maintenance of standards over time gets very difficult.

The GCSE system is complex. A system of this complexity is difficult for schools and colleges to understand, and it increases the risk of problems such as those we have seen this year. Modularisation creates a particular difficulty with maintaining standards in graded qualifications. We have already announced that we will be moving away from a modular system for GCSEs in England after the forthcoming school year. We believe that this year's experience shows that this was the right decision for students in England.

Students starting new English GCSE courses this autumn will not be doing modular exams. They will be preparing for exams and controlled assessment at the end of their course in June 2014.

Students who are part way through GCSE English courses should be reassured that when exam boards set grade boundaries in future they will be able to do so based on a more complete understanding of how the qualification works and data from this summer. We will be monitoring the delivery of this qualification carefully. It is important that schools know what conclusions it is safe to draw – and not draw – from the information they receive from exam boards and elsewhere. Exam boards will review the advice and guidance given to schools about GCSE English, including; its structure, how grade boundaries are set and how grade boundaries may be used with caution in predicting student outcomes.

As for all GCSEs and A levels, schools can make enquires about results (known as EARs) to have the marking of an individual script checked. No doubt some schools will want to do this and many have already asked their exam boards to do this. EARs will not result in changes to grade boundaries. So some schools may wish to discuss with their exam boards whether their EARs are appropriate and whether there is further information the exam board can provide about performance at the school level to guide teachers. We know exam boards are ready to do this.

Any student disappointed with the outcome of his or her GCSEs has an option to do a resit. The next opportunity for this is in January 2013. Given the concerns of some schools and colleges about their students' qualification outcomes from June 2012, each of the exam boards offering GCSE English and English Language qualifications will provide an exceptional, one-off, re-sit opportunity in November 2012. The January 2013 and June 2013 assessment opportunities for these specifications will

go ahead as planned. To do this we will need to relax some of the rules around resits. We would welcome views on the proposal by 7th September. Exam boards will confirm arrangements in the week beginning 10th September.

We set out in this report our decisions. These decisions apply to England only. There are separate issues in Wales and Northern Ireland which their governments and the regulators are considering, and we will discuss with them the conclusions that we have reached and the work they are doing.

Appendix 1: Background to the 2012 results

The 2012 GCSE results for England, Wales and Northern Ireland were announced by exam boards on 23rd August. Across the 47 GCSE subjects and almost 300 different specifications in use, the percentage of candidates achieving A*–C grades fell slightly, from 69.8 per cent in 2011 to 69.4 per cent this year. Results were broadly on a par with the results in 2010.

We set out an overview of GCSE and specifically GCSE English results for recent years in Appendix 2 to this report. In 2012, there were slight variations in GCSE results, subject by subject, as might be expected, but achievement fell noticeably in two subject areas: science, and the English subjects (English, English literature and English language), known as the “English suite”. If we look at GCSE results without taking into account science and the English suite, then results rose by 0.5 percentage points. But the changes in results for science and for the English suite were noticeable, and pulled down the aggregated figure, so that results fell by 0.4 percentage points overall.

Schools, colleges and students were generally expecting changes to achievement levels in GCSE science, and so were the regulators and exam boards. GCSE Science specifications were strengthened in response to the regulators’ concerns about the demand of the previous qualifications¹. That means that standards were changed purposefully and quite properly. Teachers were aware of the change and the reasons for it.

Teachers have been teaching the new science specifications in schools since September 2010. This is the first full award for them, and achievement at A*–C has fallen by 2.2 percentage points, from 62.9 per cent last year to 60.7 per cent this year, in line with expectations. We are not experiencing unusual levels of inquiry about GCSE Science results, or indeed about results in any subject other than those in the English suite.

In the English suite, regulators were not necessarily expecting outcomes to be exactly the same, year-on-year. This is not because standards have changed, as they have in science – standards have stayed the same. But there were a number of different factors – most especially, changes to the cohort, to the specifications and to the combinations of subjects and units taken by candidates – that could lead to different outcomes this year. A*–C achievement in English and English language

¹ See www.ofqual.gov.uk/news-and-announcements/130/336 [accessed on 31st August 2012]

(combined) fell by 1.5 percentage points. Achievement in English literature fell by 2.1 percentage points.

In particular, a new set of GCSEs has been introduced – English, English language and English literature. The expectation is that students in England would choose either English, or English language and English literature.

Also, the qualifications from all exam boards were modularised for the first time. Assessments can be taken and awarded at different times during the course, and the final grade depends on achievement across all the units. However, at least 40 per cent of the assessment must be taken at the end of the course.

The new qualifications criteria for the English suite provided for what is known as controlled assessment to replace the coursework that was part of the old qualifications. Coursework was done by candidates during the school year, and there were legitimate concerns about whether in every case coursework was genuinely the candidate's own work. Controlled assessment makes sure that coursework is done by candidates in controlled conditions, in their schools and colleges.

Qualifications in the English suite must be made up of 60 per cent controlled assessment and 40 per cent external assessment. The external assessment can be taken at one of two levels – the foundation tier and the higher tier.

For controlled assessment leading to a written outcome, awarding organisations receive a sample of teacher-marked candidate work from each school and college. For the speaking and listening controlled assessment no written outcome is produced. Moderation of the teacher's marking is via a visiting moderator from the exam board who will observe candidates completing assessment tasks and evaluate the standard of the teacher's marking.

There are issues with controlled assessment. We reported on early teething problems in October 2011, two years after it was introduced². At that stage we were concerned about the complexity of the delivery arrangements, and exam boards have since streamlined arrangements to make them more consistent between exam boards, and also easier to understand. More fundamentally, we launched a second review in April this year³ in response to what we were hearing from schools about the impact of controlled assessment on teaching and learning.

² www.ofqual.gov.uk/downloads/category/136-other-research?download=1164%3Aevaluation-of-the-introduction-of-controlled-assessment [accessed on 31st August 2012]

³ www.ofqual.gov.uk/news-and-announcements/83-news-and-announcements-news/885-changes-to-seminars [accessed on 31st August 2012]

Subject by subject, schools and colleges generally predict or estimate GCSE results for their students as a group and for their students individually. This year, in some schools and colleges, the English suite results are much better than they expected. In others, they are much worse than they expected. The impact has been noticeably different, school by school, when they compare their results with what they were expecting. This is not comparing this year's results with what was achieved in each school or college last year, or in earlier years. We believe there is a much better reconciliation there – although we need to explore this further – but even so, some schools have seen year-on-year changes that they cannot comprehend.

Exam boards are obliged to maintain standards, to make sure that, each year, students get the results they deserve. This involves making judgements on grade boundaries – on what mark in an exam constitutes a C grade, for example. Grade boundaries are different for different units within qualifications and are different year-on-year as well. This reflects the fact that different exams have different questions and are therefore at different levels of demand; a paper that is harder will need lower grade boundaries⁴. That is the nature of things, so that standards are maintained even though other things change, and schools and colleges understand that.

There are a number of units – including, but not only, controlled assessment⁵ units – where the grade boundaries set in June 2012 were higher than those set when the units were assessed earlier in the course. In some controlled assessments, this was the same unit; that is, the same task undertaken by students under controlled conditions.

In many cases, we know that teachers used their knowledge of the grade boundaries in previous units to make estimates of likely student performance in their GCSEs in June 2012. Because of the changes, those estimates proved wrong, which meant many students did not get the grades they expected.

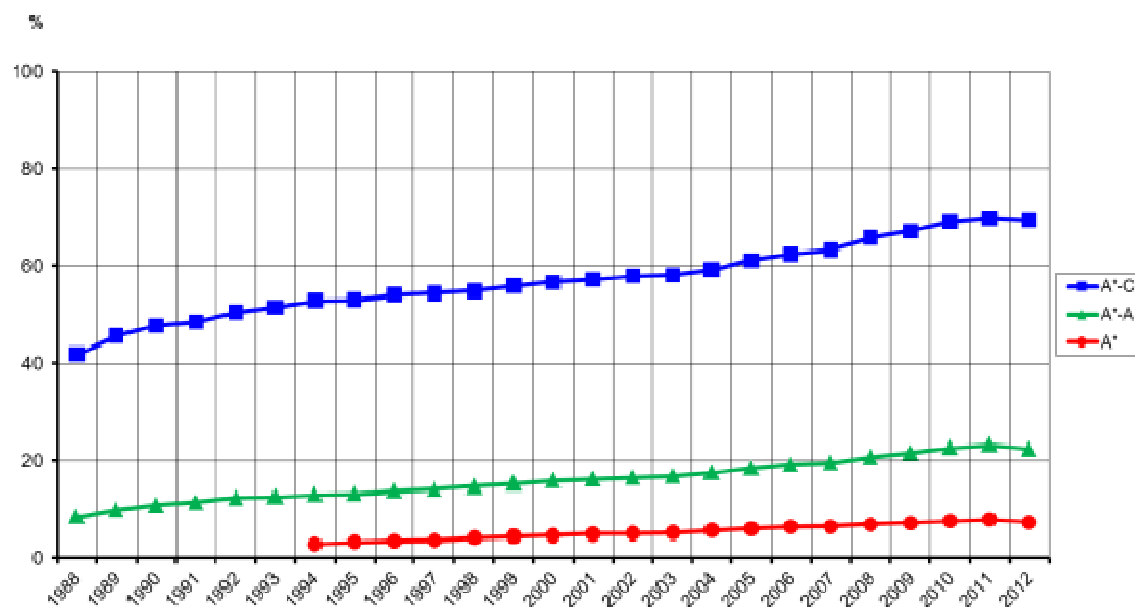
In Appendix 3 there is a more detailed explanation of how we maintain standards.

⁴ www.ofqual.gov.uk/files/2007-comparability-exam-standards-c-chapter1.pdf [accessed on 31st August 2012]

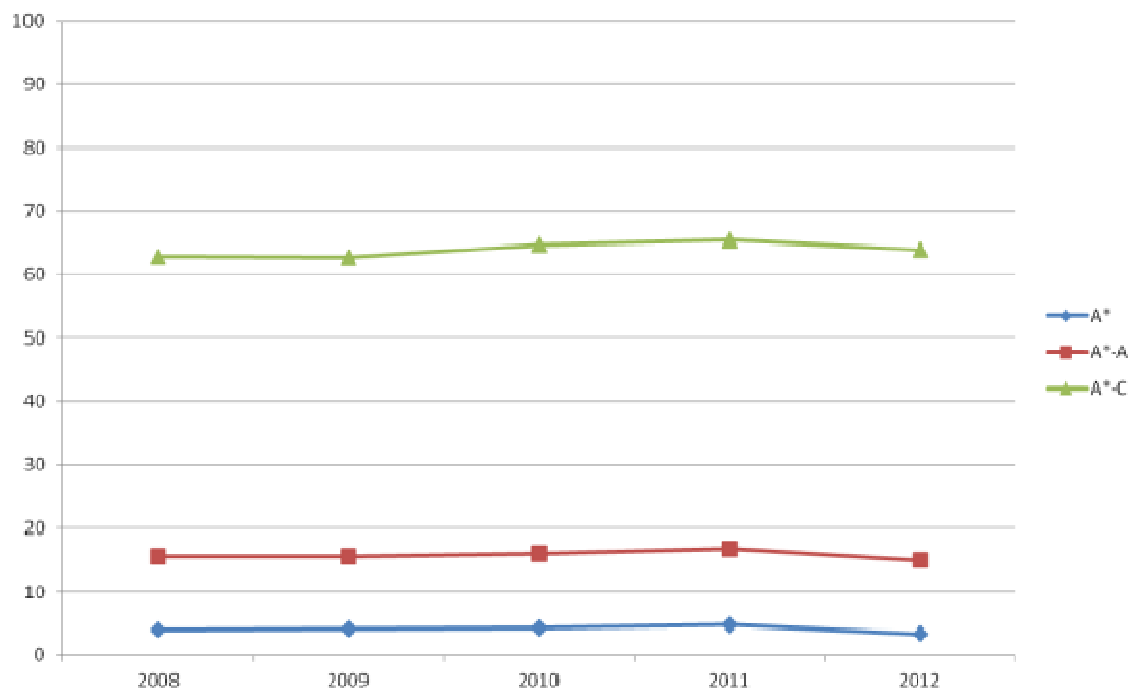
⁵ Controlled assessments are supervised and marked by teachers. They include both written work and an assessment in speaking and listening. Exam boards moderate the marks given to the assessments by teachers, using a sample of assessments. If an exam board finds a teacher has not given the work an appropriate mark it can change the marks.

Appendix 2: Summary of GCSE results over time

GCSE grades 1988–2012



GCSE English grades 2008–2012



GCSE English grades 2010–2012

Year	Total Cohort	% of candidates achieving grade					
		A*	A	B	C	D	A*-C
2010	709,831	4.4%	11.6%	20.0%	28.7%	19.7%	64.7%
2011	649,047	4.7%	12.0%	20.1%	28.3%	20.1%	65.1%
2012	674,202	3.3%	11.5%	20.2%	28.1%	20.6%	63.0%

Year	Total Cohort	Number of candidates achieving grade					
		A*	A	B	C	D	A*-C
2010	709,831	31,205	82,254	142,037	203,523	139,984	459,019
2011	649,047	30,407	78,039	130,730	183,485	130,326	422,661
2012	674,202	22,353	77,408	135,919	189,128	138,836	424,808

Figures are based on JCQ data, covering GCSE English specifications for 2010 and 2011 and GCSE English or English Language specifications for 2012, in England, Wales and Northern Ireland

All figures for 2012 GCSE results are published on the Joint Council for Qualifications website at:

www.jcq.org.uk/attachments/published/1727/GCSE%20Results%202012.pdf

[accessed on 31st August 2012]

Appendix 3: Maintaining standards – details of this approach

Awarding – what is it?

Awarding starts with the development of exam papers and assessment tasks. Examiners try to write question papers and mark schemes that are of similar demand in each exam series and year by year. The exams have to be different each time, of course, to avoid candidates knowing what they will be tested on – but each year's candidates should face equally demanding tasks. However, assessment experts accept that it is not possible to know the exact difficulty of the papers for candidates until they have answered the questions and had them marked.

A paper may turn out to be slightly more difficult or slightly easier than previous ones. The demand of the paper may be different and/or the marking may be more severe or more generous. This means that although the same mark might be given for different exams, those marks might be translated into different grades.

The lowest mark which will be awarded a particular grade on each exam is known as the “grade boundary”. Because the same mark may represent a different standard in different exams, grade boundaries must be reviewed and if necessary adjusted for each series. Awarding is the process by which this is done. Similar arrangements apply for controlled assessments.

How grade boundaries are set

The decision of where to set grade boundaries is made using all the information awarders have available to them at the time.

If the awarders set the A boundary at 70, C at 54 and F at 33, the boundary for B would be 62 – half way between the A and C boundary. The boundary for A* would be 78 and for D and E it would be 47 and 40 respectively.

Table 1: example grade boundaries

Grade	A*	A	B	C	D	E	F
Mark	78	70	62	54	47	40	33

Having agreed the grade boundary marks for the raw marks (the marks the candidates actually scored on the paper) those raw marks are converted into uniform marks (UMS marks). Those UMS marks for each unit are added together to give the overall grade.

There are particular challenges with setting grade boundaries and maintaining standards in modular qualifications. The standard needs to be maintained at the qualification level. But awarding decisions need to be made at the level of individual units, before full information is available about the performance of all candidates.

In modular schemes, relatively few candidates normally enter units early in the course, and these candidates might not be typical of all candidates who will take the units at later dates. Statistical information available at the awarding meeting held by the exam board is limited. Where a qualification is new, it is also more difficult for awarders to make judgements about the quality of work that candidates have produced in response to a new style of question paper. That makes it hard for the awarders to be certain exactly where to put the grade boundary for a unit so that it will reflect qualification level standards. In later awards much more information will be available to assist the awarders in their judgements.

When qualification awards come to be made, awarders have to work with the unit awards that have already been made. If those unit awards have not been made at the right level, then the exam boards need to review units which have not yet been awarded to try and secure standards at qualification level.

How standards are maintained

The comparable outcomes approach to maintaining standards was applied in 2009 to AS levels, in 2010 to A levels, and in 2011–12 to the new GCSEs (which were introduced over two years). We agreed our comparable outcomes approach for GCSEs in 2010. It has been the key factor in the stabilisation of results over time.

Full details of the comparable outcomes approach are set out on our website⁶. We discuss this approach regularly with our fellow regulators and with assessment experts and we keep it under review⁷. So far, experts generally agree that it is the “best practice” approach – the best way, currently, for standards to be maintained when qualifications change. We review and evaluate the approach regularly⁸. No system is perfect, but there is no known better approach to maintaining standards at the moment.

⁶ See www.ofqual.gov.uk/files/2012-05-09-maintaining-standards-in-summer-2012.pdf [accessed on 31st August 2012]

⁷ A review of our approach and how it has played out in summer awarding is set for discussion at the September meeting of our Standards Advisory Group. Membership of the Standards Advisory Group is set out in Appendix 5.

⁸ See, for example www.ofqual.gov.uk/files/2011-09-29-investigating-the-relationship-between-a-level-results-and-prior-attainment-at-gcse.pdf [accessed 31st August 2012].

There are particularly good reasons to aim for comparable outcomes in the first year of a new syllabus. Maintaining grade standards is most difficult when qualifications change. Teachers and students may have fewer resources and will have to rely on sample papers rather than actual past papers. Syllabuses may include new topics and teachers will have no direct experience of what is expected in the actual exams. Regulators and exam boards need ways in which they can be as sure as possible that when things change, standards are nevertheless maintained and students get the results they deserve.

Students taking their A levels or GCSEs in any particular year will be competing with those from other years for access to further and higher education and employment. Students have an undeserved advantage if they get better results simply because they were taking a familiar qualification. Our aim is that students should not be advantaged or disadvantaged simply because they were the first to sit a new set of exams. The only exception to this would be when a new syllabus is designed to be at a different standard – as with GCSE Science.

The principle applied in setting standards for a new GCSE is that a student should get the same grade as they would have done if they had taken the old version. We call this approach “comparable outcomes”. It aims to prevent “grade inflation” – that is, increases in the proportion of students achieving higher grades without evidence of real improvements in performance. It also adjusts for the dip in performance that can arise when a new qualification is first taken. It is more sophisticated than “norm referencing” where fixed proportions of candidates would be awarded a grade without reference to candidate achievement.

Before settling the grade boundaries, examiners use data on the prior attainments of the cohort of students to check the outcomes. This enables examiners to consider whether, for example, the awards are too generous, or whether they have not allowed sufficiently for the introduction of unfamiliar exam features. We expect exam boards to set standards within a certain tolerance, and we will challenge them if they are outside that tolerance.

The comparable outcomes approach should lead to just that – comparable outcomes, year-on-year, all other things being equal. As set out above, things are not always equal; there are sometimes changes to the cohort, for example. That is why we would expect to see small differences from year to year.

In modular qualifications, the number of candidates normally entering units early in the course is small, and these candidates might not be typical of all those who will take the units at later dates. Statistical information available at the awarding meeting held by the exam board is therefore limited. It is also more difficult for awarders to make judgements about the quality of work that candidates have produced in response to a new style of question paper. That makes it hard for the awarders to be

certain exactly where to put the grade boundary for a unit so that it will reflect qualification level standards. In later awards much more information will be available to assist the awarders in their judgements.

A modular system which includes controlled assessment together with changes to the cohort and frequent changes to qualifications themselves all make awarding a demanding business. Maintaining standards and ensuring students get the grades they deserve becomes more difficult than we would wish. 2012/13 will be the last year in which GCSE qualifications in England have modules – after that, all assessment will be at the end of the course. Then the system will be simpler and it will be easier to maintain standards.

The role of the regulator

We were established in 2008, becoming fully operational in 2010, and we regulate school qualifications in England. Arrangements are different in Wales and in Northern Ireland. (We work with the regulatory bodies in Wales and Northern Ireland with the aim of securing comparable standards in school qualifications across the three countries.)

Unlike our predecessor body, the Qualifications and Curriculum Authority, we are accountable to Parliament, not to ministers. When our enabling legislation was going through Parliament there was a political consensus that it was important to separate ministers from decisions about standard-setting, as standards measure the success of students and of the wider education system. It was agreed that being independent in this way would make it easier for us to take decisions that would secure the rigour of the system, and that over time, public confidence in qualifications – which had been damaged by allegations of dumbing down – could be improved.

Ministers rightly take an interest in our work, which has a significant impact on the wider education system. But ministers have not attempted to influence our decisions on standards, results, grades or grade setting – and if they tried to do so, we would say publicly that they had done so. Their discussions with us have been appropriate and transparent.

We were originally given an objective to secure qualification standards – to make sure that qualifications give a reliable indication of knowledge, skills and understanding and indicate a consistent level of attainment over time. The objective was amended last year to require us to consider additionally whether standards need to be raised to bring them into line with qualifications awarded overseas. We have to use our powers to try to secure this, and our other objectives. Our powers allow us to set requirements (“conditions”) which exam boards must meet, and to intervene if they do not do so. For example, we can direct them, and the direction is enforceable in the courts.

If we are to achieve our standards objective we need to have the power to intervene if an exam board is proposing to do something which does not secure standards. This may include directing an exam board to change its grade boundaries if we think that is right in order to secure standards. We would, of course, rather not have to direct: we would rather that exam boards made the right decisions themselves. But when Parliament debated our establishment, they were clear that this was a power we should have. As yet, we have not had to direct in this way.

We also have specific equality duties that we have to take account of. As a public body, we are subject to the public sector duty under section 149 of the Equality Act 2010. We have published our Equality Objectives 2012–13⁹.

⁹ www.ofqual.gov.uk/downloads/category/139-information?download=1367%3Aequality-objectives-2012-2013 [accessed 31st August 2012]

Appendix 4: Summary of meetings and actions over the last week

As part of our work we considered a wide range of information and spoke to many people from different organisations. We required each exam board to provide us with specific data about GCSE English and GCSE English language qualifications in the January and June 2011 and 2012 series. We talked to schools and teacher associations to understand their views and concerns. All of this information has informed our conclusions as set out in this report.

Data requests to exam boards

We are grateful to each exam board for providing us with the following information:

1. For each unit, for each tier (where appropriate), for each series in 2011 and 2012, the entry, the maximum mark, the grade C/D boundary mark and the proportion of candidates scoring at or above that boundary mark.
2. Copies or transcripts of the reports principal examiners presented to awarding meetings and any additional analysis made available to those meetings the recommendations from these meetings and any concerns raised by examiners in reaching the recommendations.
3. Details of any changes made to grade boundaries between the recommendations of awarding meetings and those signed off by responsible officers.
4. Analysis of the prior attainment of those certificating this summer (using either Key Stage 2 results or other methods, as appropriate) who were at or just above the C/D boundary for (a) the foundation tier and (b) the higher tier, how this compares with predictions and how this compares with performance on externally assessed papers taken in earlier series.
5. Definitive data for the qualification as a whole showing:
 - (a) for all candidates, total entry and proportions in each grade from each tier and overall in the summer 2012 series;
 - (b) for matched candidates, numbers matched and for grades A*, A, C and F, predicted cumulative percentages of candidates against actuals.
6. Identification of the particular schools with a substantial entry where the change in results from last summer has been dramatic (at least a 15 per cent change in the proportion awarded a grade C or better) and then an analysis of how this relates to entry in summer 2012/earlier and which tier paper was taken.

7. Numbers and grade distributions of any qualification awards for these new specifications made before June 2012.
8. Analysis of when candidates were entered for each unit: those entering all units in the same series, compared with those entering units in more than one series, and how this varies across the different units and series.
9. Analysis of the pattern of re-sits by candidates who certificated this summer: for each unit, the number of candidates re-sitting units – the series when they were first entered for the unit and the series when the re-sit occurred, and the distribution of grades in the unit (for example, a D in the original entry was converted to a C at the re-sit).
10. Where grade boundary changes from series to series were material, why this change was considered necessary.
11. An assessment of the accuracy of teachers' marking, based on moderation of controlled assessment units. And whether there is any evidence to indicate whether teachers' marking was influenced by the grade boundaries that had been set in the previous series.
12. Awarding committee judgements (known as the tick charts) for all units in English and English language for the January and June 2012 series.

Investigation meetings

As part of our investigation we held meetings with the organisations listed below. All meetings listed were attended by either our Chief Regulator or a Director.

Schools and school representative groups

Belle Vue Girls' School

Boston Spa School

Buttershaw Business and Enterprise College

Carr Manor High School

Dixons Academies

Horsforth School

Independent Academies Association

Landau Forte Charitable Trust

The Morley Academy and The Farnley Academy

Woodkirk Academy

Teacher associations

Association of School and College Leaders

NAHT The Association for All School Leaders

Other groups

Leeds City Council

Exam boards

AQA

CCEA

Edexcel

OCR

WJEC

We also met the Wales regulator (the Welsh Government) and the Northern Ireland regulator (CCEA).

Appendix 5: Membership of the Ofqual Standards Advisory Group

Jo-Anne Baird, University of Oxford

Paul Black, King's College London

Robert Coe, Durham University

Mike Cresswell, University of Bristol

Jannette Elwood, Queen's University Belfast

Tina Isaacs, Institute of Education

Jerry Jarvis

Sue Kirkham, Association of School and College Leaders

Barnaby Lenon, Independent Schools Council and Ofqual Board Member

Sarah Maughan, National Foundation for Educational Research

Michelle Meadows, Assessment and Qualifications Alliance (AQA)

Roger Murphy, The University of Nottingham

Paul Newton, Cambridge Assessment

Alastair Pollitt, Cambridge Exam Research

Amanda Spielman, Chair, Ofqual

Glenys Stacey, Chief Regulator, Ofqual

Gordon Stobart, Institute of Education

Peter Tymms, Durham University

Julius Weinberg, Kingston University and Ofqual Board Member

Alison Wolf, King's College London

Appendix 6: Glossary

Awarding – Where the marks candidates have been given for all units/components of a qualification are converted into overall grades.

Awarding committee – The group of people responsible for setting the grade boundaries for an exam or controlled assessment. The group comprises a chair of examiners, chief examiner and one or more principal examiners. The committee will also review overall grade outcomes for a qualification.

Cashing in – The process of claiming for certification – indicating that a candidate wishes to submit the results for a unit towards the award of a qualification. Sufficient units must be cashed in in order for the exam board to aggregate the unit results and make an award. At least 40 per cent of a GCSE assessment must be taken in the exam series in which the qualification is cashed-in.

Centre – An organisation (such as a school or college) undertaking the delivery of an assessment to candidates on behalf of an awarding organisation.

Chair of examiners – An individual responsible to the awarding organisation for maintaining standards across different specifications in a subject within a qualification and from year to year.

Chief examiner – An individual responsible to the chair of examiners for ensuring that the exam as a whole – including both internal and external assessment – meets the requirements of the specification and maintains standards from one year to the next. The chief examiner also acts as a principal examiner or moderator for at least one component.

Comparable outcomes – This is explained in Appendix 3.

Controlled assessment – Coursework carried out by students under controlled conditions in their school or college (introduced in 2009).

Exam board – Also referred to as an awarding organisation or an awarding body. An organisation recognised by us for the purpose of awarding regulated qualifications. We generally use the term ‘exam board’ to refer to awarding organisations offering GCSEs and A levels.

Examiners – Individuals with subject expertise who are responsible for marking candidates’ responses.

Foundation tier – For some subjects, all candidates will sit the same exam, but in other GCSE subjects there are two tiers: “higher” or “foundation”. Each tier leads to a

different range of grades: higher tier exams lead to grades A*–D, foundation tier exams lead to grades C–G.

GCSE (General Certificate of Secondary Education) – Generally taken as a two-year course by students aged 14–16. GCSEs sit in the National Qualifications Framework (NQF) in England, Wales and Northern Ireland. GCSEs are graded A* to G, grades D–G being at level 1 and grades A*–C being at level 2 in the NQF.

Grade boundary – the minimum uniform mark that a candidate needs to score to achieve a particular grade for an exam or controlled assessment

Higher tier – See Foundation tier.

Mark scheme – A scheme detailing how credit is to be awarded in relation to a particular assessment unit or component; a mark scheme normally characterises acceptable answers to questions or tasks or parts of questions or tasks and identifies the amount of credit each attracts.

Moderation – The process through which the marking of assessments by schools or colleges is monitored to make sure it meets required standards and through which adjustments to results are made, where required, to ensure that results are based on the required standard.

Modular GCSE – Modular GCSEs are broken down into units which are tested individually throughout a GCSE course. They are also known as unitised GCSEs. Breaking the course into modules means that candidates are able to sit assessments throughout the course. However, candidates must be entered for at least 40 per cent of the assessment for a GCSE subject in their final sitting.

Principal Examiner – An individual responsible for the setting of the question paper or task and the standardising of its marking.

Raw marks – What candidates actually scored for a unit/component. The raw marks are converted into uniform marks for use in awarding.

Responsible Officer – The person in each awarding body who is ultimately responsible for the standards of all exams offered by that awarding body, as required by our General Conditions of Recognition.

Tolerance limits – Part of the comparable outcomes approach to standard setting. Tolerance limits are set around predicted qualification outcomes to take into account the number of entries and other statistical factors for particular subjects. The tolerance limits act as triggers for exam boards to report differences from expectations.

Uniform marks – Uniform marks are given using the Uniform Mark Scale (UMS). The UMS balances out differences between exams/controlled assessments which are used for different sittings of the same unit.

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

First published by the Office of Qualifications and Examinations Regulation in 2012

© Crown copyright 2012

You may re-use this publication (not including logos) free of charge in any format or medium, under the terms of the [Open Government Licence](#). To view this licence, visit [The National Archives](#); or write to the Information Policy Team, The National Archives, Kew, Richmond, Surrey, TW9 4DU; or email: psi@nationalarchives.gsi.gov.uk

This publication is also available on our website at www.ofqual.gov.uk

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

Office of Qualifications and Examinations Regulation	
Spring Place	2nd Floor
Coventry Business Park	Glendinning House
Herald Avenue	6 Murray Street
Coventry CV5 6UB	Belfast BT1 6DN

Telephone 0300 303 3344

Textphone 0300 303 3345

Helpline 0300 303 3346