GCSEs and A Levels in Summer 2012

Our approach to setting and maintaining standards

This is an updated version of a document published in May 2011 ahead of the summer 2011 awards.

Overview

Since 2008 there have been new versions of almost all GCSEs and A levels. New A levels were awarded for the first time in summer 2010, including the new A* grade. New GCSEs in most subjects were awarded for the first time in summer 2011. There are also new GCSEs in different subjects being awarded for the first time in summer 2012 and summer 2013.

In general, the principle we have applied in setting standards for new qualifications is that a student should get the same grade as they would have done had they entered the old version of the qualification. We call this approach 'comparable outcomes'. It aims to prevent what is sometimes called 'grade inflation' – that is, increases in the numbers of students achieving higher grades where there is not sufficient evidence of real improvements in performance. It also enables us to allow for the dip in performance that can arise when a new qualification is first taken.

But in some cases we do not set out to maintain standards, because we want to set new standards – in particular if we have required changes.

This paper explains our approach and the reasons for it. It also includes background information on how awarding works now (Appendix A) and how we judge whether exam boards have followed the agreed approach (Appendices B and C).

The key points are:

- In summer 2012 we will continue to prioritise comparable outcomes in A levels, to avoid grade inflation
- In new GCSEs awarded for the second time in summer 2012, we will continue to prioritise comparable outcomes, to avoid grade inflation

- In new GCSEs in English, English language, English literature, ICT and mathematics, awarded for the first time in summer 2012, we will prioritise comparable outcomes with the previous syllabuses

- In new GCSEs in science, awarded for the first time in 2012, we will not expect comparable outcomes, because the performance standard required for the grades, particularly the higher grades, will be higher. This is because of our concerns, first reported in 2009, that the demand of the previous qualifications was not sufficient.

Maintaining standards when qualifications change

Maintaining standards is most difficult when qualifications change. Teachers and students may have fewer resources and will have to rely on sample papers rather than past papers. Syllabuses may include new topics and teachers will have no experience of what is expected in the examinations. It is also more difficult for senior examiners to make judgements about the quality of work that candidates have produced in response to a new-style question paper.

In 2001/02, discussions between exam boards and regulators about maintaining standards in the new 'Curriculum 2000' AS/A levels led to an agreement to prioritise "comparable outcomes":

> The comparable outcomes perspective implies that grade boundaries should be fixed so as to take account of any deficits in … examination performance which are unique to the first cohort of candidates. On the other hand, the comparable performance perspective entails an acceptance that candidates' results in [the first year of a new syllabus] should suffer because for this reason they did not produce performances comparable to those which would have been achieved by candidates [in the previous year].
>
> Cresswell, M J (2003) Heaps, prototypes and ethics: the consequences of using judgements of student performance to set examination standards in a time of change. University of London Institute of Education

By 'comparable outcomes' we mean the proportion of students achieving each grade. So if we aim for comparable outcomes, roughly the same proportion of students will achieve each grade as in the previous year.

When we talk about 'comparable performance' we mean the demand of the knowledge, skills and understanding that students must show in the exam. So if we aim for comparable performance in the first year of a new syllabus, it will make it harder for students to get each grade compared with students the previous year.

So there are good reasons to aim for comparable outcomes in the first year of a new syllabus. Students taking their A levels or GCSEs in any particular year will be competing with those from other years for access to higher education and employment. It gives some students an undeserved advantage if they get better results simply because they were taking an exam that their teachers were used to preparing them for. Students should not be advantaged or disadvantaged simply because they were the first to sit a new set of examinations. The only exception to this would be when the new syllabuses were designed to be at a different standard.

But what about when qualifications haven't changed? There are also good reasons to prioritise comparable outcomes after the first year of a new syllabus. If we accept that performance in the first year's exams will be a little lower than in subsequent years, and then we use the performance in those first year's exams as a benchmark for the subsequent years, we will be locking in a lower standard. Over time this is likely to result in grade inflation with, each year, gradually more candidates achieving each grade.

But we don't always want to aim for a comparable outcomes approach. We only want comparable outcomes when:

- the cohort for the subject is similar, in terms of ability, to previous years
- the syllabus and the exams and other assessments are fit for purpose
- the purpose, requirements and nature of the qualification is the same
- there has been no substantial improvement (or drop) in teaching and learning at a national level
- previous grade standards were appropriate.

Where this isn't the case, we won't always want to aim for comparable outcomes.

Maintaining standards in A levels

We followed the comparable outcomes approach for the first awards of the revised A levels in 2010. Since then we have seen much smaller increases in the proportions of students achieving top grades. The table below shows the proportions of students achieving grades A*, A and E in 2008, 2009, 2010 and 2011. The increases since 2009 are much smaller than the increases between 2008 and 2009.

A level	2008	2009	2010	2011
Grade A*	-	-	8.1	8.2
Grade A (cumulative %)	25.9	26.7	27.0	27.0
Grade E (cumulative %)	97.2	97.5	97.6	97.8

Source: JCQ provisional data for all subjects

Therefore, to avoid grade inflation in 2012 and beyond, we've agreed with exam boards that they will continue to prioritise comparable outcomes. Appendix B provides more detail of how we judge whether outcomes are comparable.

Maintaining standards in GCSEs

Revised syllabuses have been introduced in three phases, as shown below. The first phase of these new GCSEs was awarded for the first time in summer 2011.

	First teaching	First (full course) subject awards
Phase 1 – all subjects except those below	September 2009	Summer 2011
Phase 2 – English, English language, English literature, mathematics, ICT	September 2010	Summer 2012
Phase 3 – science suite (science[1] additional science, additional applied science, biology, chemistry, physics)	September 2011	Summer 2013

[1] New GCSE science will be available for certification in summer 2012

Prior to these revisions, most GCSEs were linear, so all assessment was taken at the end of a two-year course. The new syllabuses are almost all unitised. Students may re-sit each unit once but they must also sit at least 40% of the assessment at the end of the course. As of September 2012, we will once more be requiring GCSEs taken in England to be linear.

In 2010 we agreed with the exam boards that summer 2011 outcomes should be comparable with those in summer 2009[2]. It was also agreed that standards at unit level should be consistent with the previous syllabuses and would not take account of students' relative immaturity when entering units early.

Since 2011 was the last year of the old ('legacy') syllabuses for most subjects, we agreed that exam boards would continue their previous approaches to maintaining standards in those subjects.

The table below shows the summer 2011 results for the new subjects (phase 1 above) and legacy subjects (phases 2 and 3 above).

GCSE	2010	2011
Grade A* - new subjects	9.2	9.0
Grade A* - legacy subjects	6.5	7.1
Grade A (cumulative %) – new subjects	26.2	25.8
Grade A (cumulative %) – legacy subjects	19.0	20.4
Grade C (cumulative %) – new subjects	70.7	70.4
Grade C (cumulative %) – legacy subjects	68.1	69.7

Source: JCQ provisional data for all subjects

The outcomes for the legacy subjects (shaded boxes) increased from 2010 to 2011. One of the reasons for this is likely to be that teachers are used to the syllabus and more familiar with what their students need to do to achieve high grades.

[2] It was agreed that 2009 would be used for comparison as this was the last year when only the old syllabuses were available. In summer 2010, both the old syllabuses and units for the new syllabuses were available for students to enter.

In contrast, the outcomes for the new subjects, where we were aiming for comparable outcomes, were consistent between 2010 and 2011.

We will continue to aim for comparable outcomes in the phase 1 GCSEs in summer 2012, to avoid grade inflation. We will also be aiming for comparable outcomes in the phase 2 subjects – English, English language, English literature, ICT and mathematics.

In February 2012 we announced moves to strengthen GCSEs in English literature, geography, history and mathematics. New syllabuses for geography will be taught from September 2012 and new syllabuses for English literature and history will be taught from September 2013. Mathematics papers will be improved from November 2012 onwards. All exams in English literature, geography, history and religious studies will include separate assessment of spelling, punctuation and grammar from September 2012 onwards.

GCSE science suite

New syllabuses for the GCSE sciences were introduced for teaching from September 2011. The science suite includes the following GCSEs:

- Science
- Additional science
- Additional applied science
- Biology
- Chemistry
- Physics

GCSE science is available for certification for the first time in summer 2012. The other GCSE science subjects in are available for certification for the first time in summer 2013. When we refer in this document to 'GCSE science' we mean that particular GCSE rather than the suite of science subjects.

The new syllabuses for the suite of science subjects were designed to be more challenging and to discriminate more effectively between candidates of different ability levels. This was in response to concerns reported by Ofqual in 2009[3] that the previous syllabuses didn't adequately test the subject content and weren't sufficiently demanding. This meant that students achieving higher grades didn't show the level of performance expected, as set out in the grade descriptions. To address these

[3] www.ofqual.gov.uk/files/ofqual-09-4148_GCSE_science_2007_2008_report.pdf

concerns, the new exams contain a wider range of question types, including questions which require extended writing.

Our expectation has always been that students will have to perform at a higher standard on the new exams to achieve the same grades as students in previous years, particularly at the higher grades. Therefore we are not aiming for comparable outcomes in the new GCSE science awards this summer. We do expect the grade standards to be comparable across exam boards in the new GCSE science in summer 2012, and we have been working with them to achieve this. We will require exam boards to set their grade standards so that candidates achieving the higher grades demonstrate the knowledge, skills and understanding set out in the grade descriptions. We will be reviewing the outcomes from the new GCSE science awards before results are issued in August to judge whether grade standards across all exam boards are consistent and appropriate. If necessary, we will require exam boards to change their grade boundaries. We will also review script evidence after the results are issued, to see if the candidates met the performance standards set out in the grade descriptions, and we will take action in the following year to tighten standards if necessary.

Appendix A

Background to A level and GCSE awarding

A levels were introduced in 1951, initially offered by eight different examining boards, with The Associated Examining Board (AEB) joining in 1953. Certificates showed only a pass or fail, but unofficial grades also began to be used. These varied between the boards. For example, the University of Cambridge Local Examinations Syndicate (UCLES) used a scale of 1 to 9 and the Joint Matriculation Board (JMB) reported achievement in percentages rounded to the nearest 5%.

In 1963 the Secondary Schools Examinations Council (SSEC) introduced a national A-E grading system accompanied by guidelines for the proportions of candidates within each grade band, as shown below.

Grade	Percentage of entry	Cumulative percentage
A	10	10
B	15	25
C	10	35
D	15	50
E	20	70
O*	20	90
F	10	100

* Allowed Ordinary

These norms remained in place until, in 1987, the Secondary Examinations Council (SEC) asked the boards to award A level grades on the basis of examiners' judgement of the quality of work at three key grade boundaries – A, B and E.

A levels are currently offered by five exam boards: three based in England, one based in Wales and one based in Northern Ireland. Since schools and colleges in England, Wales and Northern Ireland can enter with any one (or more) exam board, Ofqual works closely with the regulators in Wales (the Welsh Government) and Northern Ireland (CCEA) in order to ensure a common approach to A levels.

How grades are set

The principle that senior examiners use their judgement to set boundary marks for particular grades has remained largely unchanged since A levels were introduced. The following extract details the approach taken in the 1960s.

> When the chief examiners are satisfied that marking to a common standard has been achieved, the final meeting of examiners is held. They, in conjunction with the Secretary and his staff, who provide the necessary link between groups of examiners working over a wide range of subjects, proceed to consider what should be the minimum mark for a pass in the subject at Advanced level. In addition to their own views on the standard of work as seen in the scripts which they themselves have marked, the chief examiners receive detailed reports from each member of the panel. They are also provided by the Secretary with detailed statistical information about the way in which the marks in a subject as a whole run in the year under review as compared with previous years. If the distribution of marks is significantly different from that of previous years they must attempt to establish on the evidence available to them whether there has been an improvement (or deterioration) in the quality of the entry or whether the question paper, despite all the efforts made at the preparatory stage…, has proved to be easier (or more demanding) than the corresponding paper set in the previous year.
>
> Christopher (1969) JMB/GCE: the work of the Joint Matriculation Board
> (quoted in Christie, T and Forrest, GM (1980) Standards at GCE A-level: 1963 and 1973. Schools Council Research Studies. London: Macmillan Education)

The current process is very similar. Each exam board will convene an awarding committee for each subject or syllabus being awarded, chaired by the Chair of Examiners. (If there are two syllabuses offered by an exam board in a subject there will be two separate awarding committees, chaired by a single Chair of Examiners for the subject.) The awarding committee comprises the chief examiner and any principal examiners (those who write any question papers), one or more principal moderators (responsible for any coursework units), the reviser (responsible for 'revising' the question papers to ensure consistency) and an exam board officer.

The awarding committee will meet face to face or communicate remotely to consider the available evidence and make recommendations as to where the grade boundary marks should be set.

Grade boundaries are set on each unit in turn. Boundaries for two of the grades – A and E – are set in this way. These are called judgmental grade boundaries. The other boundaries – B, C and D – are calculated arithmetically. So, for example, awarders might set the A boundary at 70 and the E boundary at 38. The difference

between A and E is therefore 32 marks, so the boundaries for B, C and D would be calculated as 62, 54 and 46 respectively, as shown below.

Grade	**A**	**B**	**C**	**D**	**E**
Mark	70	62	54	46	38

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[4] marks) and those UMS marks for each unit are added together to give the overall grade.

Having considered all the units, the awarding committee will review their decisions in the light of the outcomes for the qualifications as a whole (at AS and A level), and then the Chair of Examiners will formally record their decisions.

The process for GCSE is very similar, but the judgemental grade boundaries are different – A, C and F at GCSE.

[4] Uniform mark scale

Appendix B

Predictions for A level based on prior GCSE achievement

Exam boards use the relationship between GCSE performance and A level outcomes in previous years (for example GCSE 2007 to A level 2009) to give an indication of the overall level of achievement for the cohort.

This is the same methodology that is used by systems such as ALIS[5] and ALPS[6] to predict individual student outcomes based on their GCSE results. However, there is a crucial difference in that exam boards are looking at the whole cohort (the entry they have for a particular subject) rather than at individual candidate level.

The exam boards work together to produce the predictions for A levels, based on the prior relationship between GCSE performance and A level performance. They then separate that out according to the entry that they each have, so Edexcel have predictions that relate to their entry, AQA have predictions that relate to theirs, and so on.

Using these predictions means that we can take account of differences between the entries for different exam boards. In any particular subject, if one exam board had an entry that comprised very high ability students, the predictions would suggest that the exam board would have a high proportion of students achieving grade A. Expecting each exam board to have the same proportion of candidates achieving grade A might seem to be fair but it can result in some candidates being undeservedly advantaged or disadvantaged according to the exam board they use.

The predictions also provide a common measure for reporting outcomes to the regulators in advance of results. It is important for the regulators to be able to look across exam boards, in advance of results being issued, to ensure there is a consistent approach, in the interests of fairness to candidates.

---

[5] Advanced Level Information System which provides performance indicators for post-16 students based on GCSE data (http://www.cemcentre.org/alis/introduction)

[6] **Advanced Level Performance System which provides a statistical analysis of a school or** college's AS and A level results against national benchmarks (http://www.alps-va.co.uk)

Appendix C

Predictions for GCSEs

Until 2010 different exam boards made use of different statistical evidence, including data from Key Stage 3 (KS3) national tests formerly taken in England, to predict changes in the likely results for a cohort. With data from KS3 tests no longer available as the tests stopped after 2008, in preparation for the summer 2011 exam boards sought other data to use to compare the relative ability levels of the 2009 and 2011 cohorts. We expect exam boards to use as wide a range of qualitative and statistical evidence as possible in guiding their awards.

Having considered the issues above, the regulators have agreed with exam boards that emerging results in August 2012 will be reported to the regulators using two measures. All exam boards will report their outcomes compared to the results achieved by common centres[7] from 2011. In addition, the three exam boards based in England will report their outcomes against predictions for the cohort based on prior achievement at Key Stage 2.

[7] A common centre is a centre that has entered students for a subject in the two years in question (in this case in 2012 and 2011). The assumption is that the centre's results are unlikely to be very different in those two years, and that across the cohort as a whole, comparing results for the common centres gives an indication of whether standards between years are comparable.