

Review of Standards in GCSE Chemistry

2003 and 2008



April 2012

Ofqual/12/5152

Contents

Executive summary	3
Findings	3
Section 1: Introduction	4
Context	4
Methodology	5
Provision of assessment materials and student work	5
The review team	6
Analysis of the specifications and assessment materials	7
Analysis of student performance	7
Section 2: Subject demand in GCSE chemistry	8
Overview	8
Findings	10
Assessment objectives	10
Specification content	11
Schemes of assessment	12
Assessment weightings of the extension material	13
Weightings of internal and external assessment	14
Options	15
Question papers	16
Coverage of objectives	18
Accessibility	18
Opportunities for high-attaining students	19
Coursework	20

Section 3: Standards of performance	22
Overview	22
Findings	22
Conclusions.....	23
Appendix A: Provision of assessment materials and student work at GCSE and GCE levels for Ofqual's archive (annual inclusion and standards reviews)	25
Section 1: Specification of requirements.....	25
Section 2: Student work.....	26
Appendix B: Assessment objectives	28
Appendix C: Schemes of assessment	30
Appendix D: GCSE specifications reviewed.....	35
Appendix E: GCSE scripts reviewed.....	35
Appendix F: Availability of specification materials for the purposes of this review	36
Appendix G: Student achievement by grade	37
Appendix H: Review team	38

Executive summary

The Office of Qualifications and Examinations Regulation (Ofqual) undertakes a rolling programme of reviews across high-profile GCSE and GCE A level subjects to monitor whether standards in assessment and student performance have been maintained over time.

This report details the findings for GCSE chemistry in the years 2003 and 2008. The previous review for this subject compared the years 1998 and 2003. The findings of that review were published in 2005.

The review compared subject specifications, assessment materials and student work from the five organisations awarding this qualification in the years being reviewed (AQA, CCEA, Edexcel, OCR and WJEC) by collecting the views of a number of subject specialists.

Findings

- Changes made to the nature and emphasis of the assessment objectives (AOs) – for example increasing the proportion of assessment allocated to application of knowledge and understanding (AO2) – made the qualification more demanding in 2008 than it was in 2003.
- The inclusion of How Science Works in the 2008 assessment objectives – introducing concepts such as understanding how scientific evidence is collected, analysed and evaluated in terms of validity and reliability when presenting and justifying conclusions (AO1 and AO3) – made the qualification more demanding in 2008 than it was in 2003.
- Variations in the nature of schemes of assessment, for example the relative weighting of external and internal assessment, and the styles and types of assessment instruments (such as type of question or type of task) mean that students are assessed against different combinations of the AOs, which may have an impact on the demand of the qualification experienced by different students.

Section 1: Introduction

Context

In his Review of Qualifications for 16–19 Year Olds (1996), Lord Dearing made several recommendations to ensure that “there is a basis and accepted procedure... for monitoring and safeguarding standards over time”. In the same year, the School Curriculum and Assessment Authority (SCAA), one of our predecessors, and the Office for Standards in Education, Children’s Services and Skills (Ofsted) jointly recommended that there should be:

a rolling programme of reviews on a five-year cycle to ensure examination demands and grade standards are being maintained in all major subjects. (Standards in Public Examinations 1975 to 1995, page 4, 1996)

As a result of these recommendations we, in collaboration with the Welsh Government and the regulator for Northern Ireland (Council for the Curriculum, Education and Assessment (CCEA)), introduced a programme to investigate standards in GCE A level and GCSE examinations by systematically collecting and retaining assessment materials and student work to enable standards reviews to cover two or more years.

The Apprenticeships, Skills, Children and Learning Act 2009 formalised our role in undertaking such reviews by including a statutory objective “to secure that regulated qualifications indicate a consistent level of attainment (including over time)”.

We report on our work in meeting this objective. And we use our findings to inform developments in qualification and subject criteria to support meeting this objective in the future. In our reviews we:

- analyse the nature of the requirements that different assessments make on students
- compare the levels of performance required for a particular grade in different assessments
- consider how these two elements relate to each other.

In 2003 there were 25,312 students taking the GCSE Chemistry qualifications being reviewed. In 2008 the number was 68,655. A detailed breakdown of student-entry numbers and cumulative percentage pass rates can be found in [Appendix G](#).

Our immediate predecessor, the Qualifications and Curriculum Authority (QCA), most recently conducted a standards review in GCSE chemistry, using materials from 1998 and 2003. The findings were published in a report in 2005, which is available on our website at www.ofqual.gov.uk/files/12890_chemistryreport.pdf.

Methodology

Standards reviews examine different specifications within a qualification, the associated assessment instruments and samples of student work, by collating and analysing the views of a number of subject specialists. The following sections of this report detail how we collect and process this information. In these reviews, demand is measured against that of the other specifications under review and includes consideration of:

- specification-level factors such as assessment objectives (AOs), content and structure
- assessment-level factors such as what content is assessed, the weighting of each component and how the assessments are marked
- student performance-level factors, including how the students responded to the assessments and the grades they received as a result.

The demand of an assessment or qualification can be defined in a variety of ways and is linked to the purpose of the qualification. It is related to the:

- amount and type of subject knowledge required to be assimilated
- complexity or number of processes required of the students, the extent to which the students have to generate responses to questions from their own knowledge or the extent to which resources are provided
- level of abstract thinking involved
- extent to which the students must devise a strategy for responding to the questions.

Provision of assessment materials and student work

Each of the five awarding organisations offering the qualifications being reviewed (AQA, CCEA, Edexcel, OCR and WJEC) was asked to provide specification

materials for GCSE chemistry (from the specification with its largest entry in summer 2008).

Our requirements for the provision of assessment materials and student work for review are given in [Appendix A](#) and, in summary, include:

- the current specification
- all associated question papers
- final mark schemes
- the 2008 chief examiner's report and grade boundaries, overall and by unit (both raw and scaled)
- mark distributions, grade descriptors and assessment grids
- any other information that was routinely supplied to centres
- all the assessment work carried out by a sample of students whose final grade lay at or near the judgemental grade boundaries for the qualification being analysed.

The equivalent materials that were collected and retained for the previous review were retrieved from our archive of assessment materials and student work.

Full details of the materials supplied by awarding organisations can be found in [Appendix E](#) and [Appendix F](#).

The review team

We contracted 11 experts in GCSE chemistry to undertake the review. These reviewers were sourced through:

- a subject-expert recruitment exercise carried out by us in November 2008, advertised via The Times Educational Supplement and our website and newsletter
- nominations made by awarding organisations involved in the review
- nominations made by subject associations and other learned organisations invited to participate in the review.

A full list of reviewers can be found in [Appendix H](#).

We contracted a lead reviewer, specification reviewers and script reviewers. (All nominees from awarding organisations and subject associations were script reviewers.)

Analysis of the specifications and assessment materials

The lead reviewer and specification reviewers (specification review team) analysed the awarding organisations' materials, using a series of forms which can be found via the comparability page on our website at www.ofqual.gov.uk/standards/research-reports/92-articles/23-%20comparability .

These analyses are designed to describe the demand of the specification. Each reviewer analysed a subset of the specifications available, so that there were at least three different views on each specification. The lead reviewer then produced a report which brought together the views of the reviewers on each of the awarding organisations' specifications. The specification review team was given the opportunity to discuss the lead reviewer's conclusions at a follow-up meeting. These findings are presented in Section 2 of this report.

Analysis of student performance

To assess student performance, all reviewers were brought together for a two-day meeting to analyse student scripts (pieces of student work supplied by the awarding organisations). This process is referred to as a script review. The meeting started with a briefing session to make sure that all the reviewers had a common understanding of the methodology and the judgement criteria.

The scripts were organised into packs for consideration during the review. Packs were organised by grade: A/B, C/D and F/G for GCSE. (Other grades are calculated arithmetically after the above grade-boundary marks have been set during the awarding process carried out by awarding organisations.)

Section 2: Subject demand in GCSE chemistry

Overview

Specification reviewers considered the amount and type of knowledge about chemistry required by each awarding organisation's specifications. They did this by analysing specification documents, chief examiners' reports and question papers with associated mark schemes from each of the awarding organisations in 2003 and 2008. Details of the specifications included in the review are given in [Appendix F](#).

In 2003 awarding organisations had very similar subject content in their specifications and similar schemes of assessment. However, in 2008 there was a much greater variety in awarding organisations' specifications due to the change in the nature of the AOs, different schemes of assessment and different styles of question.

For CCEA, the specification, the AOs and the scheme of assessment have not changed from 2003, and the demand of the examination papers is very similar. However, in 2008 all other awarding organisations changed their specifications and assessment models and the reviewers judged that the CCEA specification is now more demanding than a number of the others.

OCR has integrated the ideas and philosophy of How Science Works into the specification. It has addressed all aspects of the AOs within its externally set and assessed written papers. Problems with How Science Works have been specifically addressed in the Unit 1 paper and the internally set and assessed case study. However, the significant use of controlled-response (multiple-choice, objective-style) questions and the depth of response required in the externally set papers are not considered to be of sufficient demand to provide enough opportunities for students to match the appropriate grade descriptors. Overall, the specification was considered to be of a lower standard and less demanding than it was in 2003.

AQA has also integrated How Science Works into the specification. Some aspects of the AOs have not been fully addressed in the externally set and assessed papers, but the investigative skills assignment activities redress the balance.

The reviewers thought the question papers were more challenging than those in 2003, particularly for foundation-tier students.

Each unit in the Edexcel specification begins with a rationale and addresses How Science Works in broad terms, but the specification statements failed to back this up by identifying specific opportunities. In the reviewers' judgements, few questions in the written papers, C1a, C1b, C2 and C3, were devoted to the assessment of How Science Works. However, in the internally assessed activities there were questions addressing How Science Works, but the distribution was somewhat variable in the different papers that were reviewed. In practice, Edexcel has not taken the AOs and

applied them in a way which makes the Higher-tier papers sufficiently demanding. By contrast the foundation tier was comparably demanding through this period.

In 2003 the WJEC specification was large, varied and demanding. It contained academic and traditional content and was probably of less relevance to many students than it is today. However, some topics in the 2008 specification did seek to make the specification more relevant and accessible. How Science Works was not integrated into the scheme of assessment and few questions addressing this aspect were apparent in the external written papers. However, the internally assessed practical test or extended report or the investigatory planning exercise would have provided more opportunities. The length of papers in 2003 produced a challenging and demanding examination experience. The increased weighting given to AO2 in 2008, in the reviewers' judgements, made the Foundation tier more demanding for students.

AQA and OCR addressed the introduction of How Science Works in 2008 in a more detailed way in their specifications than did the other awarding organisations. They included a significant proportion of questions, in the external written papers and in the internal assessment activities and tasks, devoted to the assessment of How Science Works. The other awarding organisations gave less emphasis to these aspects certainly in the external written papers and, to a certain extent, in the various internally assessed activities and tasks. However, OCR used mostly controlled-response questions in the written papers to assess How Science Works, and these did not provide sufficient opportunities for students – particularly the higher-achieving ones – to show what they knew, understood and could do. The AQA questions were more demanding and needed students to use their scientific knowledge and understanding of How Science Works when arriving at suitable answers. The reviewers thought that the other awarding organisations did not address fully the issues of How Science Works.

After balancing the various aspects involved, the reviewers thought that AQA's externally set and assessed papers in 2008 were slightly more demanding than in 2003, CCEA's about the same, WJEC's and Edexcel's less demanding, and OCR's the least demanding. The increased range and complexity of the activities and tasks used in the internally assessed components makes it difficult to establish and judge clearly the effect on the level of demand for students. However, these components have the potential to be both accessible to weaker students and challenging for the most able.

Findings

Assessment objectives

The AOs are listed in [Appendix B](#). There was a significant change in the aims and nature of GCSE chemistry from 2003 to 2008. In 2003 and 2008 all awarding organisations followed a uniform approach, using three overarching AOs: AO1, AO2 and AO3, as defined in the GCSE criteria for science. For CCEA over this time period, the wording within each AO was identical, since the specification had not changed. However, for all other awarding organisations there was a significant change in the spirit, nature and wording of the AOs and also their relative weighting.

In 2008 AO1 became “knowledge and understanding of science and How Science Works” and there was a decrease in weighting compared with 2003. AO2 involved the application of skills, knowledge and understanding” with an increased weighting compared with 2003. In 2008 How Science Works was seen as an important aspect to integrate into teaching and learning and its consequent assessment. The importance of understanding how scientific evidence is collected, analysed and evaluated in terms of validity and reliability when presenting and justifying conclusions was highlighted (AO1 and AO3). In addition the importance of understanding how decisions about science are made in different situations was also specified (AO2).

In 2003 AO3 exclusively involved the assessment of investigative skills. In 2008 a different emphasis and a much broader range of skills were involved, including “practical, enquiry and data handling”. However, because of the variation in methods of assessment of practical skills between awarding organisations, it is difficult to identify and judge what the effect of the changes has been on the level of demand for students.

In 2003 all awarding organisations allocated weighting within the range of 45 to 55 per cent for AO1, 25 to 35 per cent for AO2 and 20 to 25 per cent for AO3, differences of between 5 and 10 per cent. In 2008 the weighting changed, and the range widened between the various awarding organisations to 20 to 55 per cent for AO1, 25 to 55 per cent for AO2 and 20 to 40 per cent for AO3.

The change in weighting in 2008, accompanied by a change in wording, was regulatory. All the awarding organisations, with the exception of CCEA, reduced the weighting of AO1 (knowledge and understanding) and increased the weighting of AO2 (application of knowledge and understanding). Edexcel had the largest range of target figures, reflecting the different assessment routes available in that specification. The increased range in the weightings raises the issue of comparability within each assessment profile.

Awarding organisations used a variety of methods and approaches to meet these requirements. Some appeared more successful than others. The introduction of How Science Works, and the increased weighting of AO2 involving the higher-order skill of application, was judged to have made the specifications more demanding particularly for foundation-tier students.

Specification content

In general, the subject content defined in all the specifications in 2003 and 2008 allocates all topic areas to both foundation and higher tiers, with particular statements within each topic designated as higher tier only.

The presentation of the specifications has changed since 2003. OCR now identifies “issues for citizens” and “questions that science may help to answer”. WJEC precedes each section of content by a short list of questions which provides the theme for the section. Edexcel includes: “Have you ever wondered?” questions to engage students with the topic content. AQA, at the beginning of each section, describes activities which develop students’ knowledge, skills and understanding of How Science Works. This is followed by the subject content students will need to develop their understanding of How Science Works. The way in which each awarding organisation has identified the requirements of How Science Works within their specifications is broadly similar but varies significantly in terms of the level of detail and development. All awarding organisations highlight the parts of the specification content which lend themselves to activities that allow How Science Works to be considered. Some specifications included How Science Works statements within the organisation of the specification content. In addition, AQA and OCR have significant separate sections devoted to all aspects of How Science Works.

To compensate for the additional teaching time involved in delivering How Science Works, the subject-content sections of the specifications have been reduced (except in CCEA’s). There is some evidence of this throughout all the awarding organisations’ specifications but to different extents. There is less industrial chemistry, fewer earth sciences and a narrower exposure to the variety of chemical reactions, but the key principles and concepts of chemistry appear to have been maintained in the specifications.

There is also evidence that awarding organisations have increased the range of topics by including more relevant and new technology:

- AQA – plant oils, emulsions and analytical methods
- WJEC – nanoscience and Smart materials
- Edexcel – designer products and nanotechnology
- OCR – liquid and gas chromatography.

However, some of these topics are considered to be relatively undemanding in chemistry terms and require recall rather than application of knowledge in examination questions at this level.

Overall, the total number of topics was probably unchanged over the time period of this review. This means that in 2008, students generally have a very similar range of chemistry to appreciate and understand as they did in 2003.

It is difficult to make an overall judgement but the introduction of How Science Works and the increased weighting of AO2 could make the qualification more demanding and put extra pressure on students, particularly at the foundation level.. However, the awarding organisations have shown different levels of effectiveness when carrying this through to the operational, externally set papers and the internal assessments.

Overall the reviewers thought that the qualifications were similarly demanding in 2008 as they were in 2003, apart from AQA's which were more demanding due to the nature and depth of the topics included.

Schemes of assessment

The schemes of assessment for GCSE are detailed in [Appendix C](#). In 2003 all awarding organisations had very similar schemes of assessment. Written papers involved structured questions and investigations of the style specified by the Scientific Enquiry (Sc1) section of the National Curriculum. Awarding organisations used one or two papers in their schemes of assessment for each tier. Furthermore, the question papers in 2003 generally tended to include at least one longer paper of 1 hour and 30 minutes or more, regardless of tier of entry.

In 2008 all awarding organisations, except CCEA, increased the number of papers to three or more within their schemes of assessment. These papers were of much shorter duration, from as little as 20 minutes to no more than 60 minutes.

So, as regards the externally set and marked papers, the number of papers taken has increased for all awarding organisations (except for CCEA), but the amount of examining time has stayed approximately the same apart from for one of the Edexcel optional routes. The change from a mostly terminal examination system to a staged-assessment approach, and the increased opportunity for students to re-take examination papers, has made the qualification less demanding for students.

The types and styles of question used in the externally written papers have also changed significantly. In 2003 all papers included structured questions with varying degrees of opportunity for "extended prose". In 2008 controlled-response questions as well as structured questions were used, and there was less opportunity for extended prose in the externally set papers. Controlled-response questions require

students to select a response from a list of possible answers or connect possible answers which are provided in other ways. Structured papers require students to generate their own responses and communicate appropriately. There was concern that the controlled-response questions did not enable the full range of students to show what they knew, understood and could do. The reviewers thought that the full range of skills required by the AOs was not being assessed by this type of question, particularly the higher-order skills. Whilst there could be a wider coverage of the specification content, the depth and demand of the assessment might be lower, partly due to the shorter length of individual papers. Furthermore, many of the controlled-response questions were worth a small number of marks and were by their nature self-contained. They did not forge a link with subsequent questions and were often lacking in progressive demand and depth. However, some of these concerns were offset, to a certain extent, by the wider and broader nature of the internally assessed components.

Assessment weightings of the extension material

Generally, there is more consistency between the awarding organisations in the assessment weighting of the extension material, with Edexcel decreasing the weighting and OCR increasing the weighting compared with 2003. Examination papers were available at each tier except, in Edexcel's case, where only one paper was used to discriminate across the full range of students' performances from A* to G grade.

As regards the assessment of How Science Works, awarding organisations used a variety of methods and approaches to meet the challenges of the change in the 2008 specifications, which included:

- questions within externally set and marked examination papers
- questions within externally set but internally assessed papers
- centre-devised, internally set activities, internally assessed and externally moderated
- a combination of these approaches.

In relation to the assessment of practical skills, in 2003 all awarding organisations included the assessment of investigative skills, using internally set activities which were marked internally against defined common criteria. In 2008 a much greater variety of approaches was adopted, assessing a range of practical skills, and skills and abilities related to How Science Works.

Therefore a range of approaches is on offer for awarding organisations to assess practical tasks, such as:

- centre-devised, internally set, internally marked activities or holistic assessment of practical skills but not moderated externally
- a practical activity followed by an externally set, internally assessed test related to the practical activity, which takes place in controlled conditions
- centre-devised assessment to assess analysis of data and applications, and implications of science
- centre-devised assessment of investigative skills, internally set, internally marked and externally moderated.

WJEC, Edexcel and OCR incorporated components which used alternative methods of assessment, allowing individual students to use and show different aptitudes and skills. Edexcel provided externally set tests which were internally marked as alternatives to centre-devised activities. WJEC included different types of practical exercise, some externally set and some internally devised. OCR included the opportunity to submit either case-study and data-analysis assessments or a practical investigation.

Weightings of internal and external assessment

Awarding organisation	2003		2008	
	External assessment (%)	Internal assessment (%)	External assessment (%)	Internal assessment (%)
AQA	80	20	75	25
CCEA	75	25	75	25
Edexcel	80	20	30–70	70–30
OCR	80	20	66.7	33.3
WJEC	80	20	75	25

For all awarding organisations, apart from CCEA, the weighting of internal assessment increased relative to external assessment. The weighting allocated to

internal assessment in the Edexcel specification can vary significantly from 30 to 70 per cent. Students will therefore have been assessed against different combinations of the AOs, and there were concerns about ensuring that grades are comparable. The large number of assessment routes available in this case is in contrast to the limited number of routes available in the specifications from the other awarding organisations.

The reviewers were concerned about the comparability and equivalence of these different types of assessment and found it difficult to make secure judgements about changes in level of demand. The provision of alternative methods of assessment, which use different contexts to assess skills and abilities, do not look comparable at first glance. However, the wording of AO2 and AO3 is such that there is greater flexibility in interpretation in terms of the type of activity which could match the wording of the AOs.

Overall, the schemes of assessment have become much broader with more variation both within and between awarding organisations. They are generally more complex than those used in 2003 and certainly allow the opportunity for a greater variety of skills to be assessed. Therefore, there is a spectrum of different opportunities.

Options

In 2003 all topics were compulsory within each of the specifications considered apart from OCR's, where there were two different routes (options A and B). In 2008 these different routes for OCR are converted into two separate specifications.

In 2008 all topics are compulsory within each specification, and the options only relate to different approaches to assessment, for example:

- multiple-choice and/or structured paper
- externally set examination papers and/or centre-devised activities
- different, centre-devised activities.

Question papers

Awarding organisation	Tier	2003 question types	2008 question types
AQA	Foundation	Structured (80%)	Structured (50-75%) and controlled-response (25-50%)
CCEA	Foundation Higher	Structured (75%)	Structured (75%)
Edexcel	Foundation Higher	Structured (80%)	Controlled-response (20-30%) and structured (10-40%)
OCR	Foundation Higher	Structured (80%)	Controlled-response (33.3%) and structured (33.3%)
WJEC	Foundation Higher	Structured (80%)	Mostly structured and a few controlled-response (75%)

For AQA, Edexcel and OCR (although depending on the route adopted by students), there is an increase in the use of the controlled-response type of question since 2003. This has decreased the opportunity for students to use extended prose in their answers in the externally set and marked papers. Although it has been offset to a certain extent by the various internally assessed components, which give more opportunities and which have a higher overall percentage weighting when compared with 2003. The reviewers thought that the “distractors” (incorrect answers in multiple-choice questions) in some of the controlled-response questions directed at How Science Works could sometimes be discounted on grounds other than students using their scientific knowledge and understanding.

There is, in general, a decrease in the percentage of the marks allocated to numerical calculations on both foundation and higher tiers from 2003 to 2008 on the externally set and marked papers. The percentage of marks allocated to writing symbol equations has not changed significantly, but considering it is an essential aspect of chemistry the percentage is regarded as being lower than it should be. However, there is more variation between awarding organisations in terms of the

percentage of marks allocated to extended prose. AQA's percentage is considerably higher than that of the other awarding organisations with Edexcel's, WJEC's and OCR's being particularly low.

The high percentage and structure of the controlled-response questions in the 2008 OCR scheme of assessment (Unit 1 and 2 papers) makes the assessment/qualification considerably less demanding in terms of the complexity of the questions and the deployment of the associated strategy skills required by students when answering them.

CCEA has an unchanged scheme of assessment and the allocation of marks for numerical calculations, balancing chemical equations and the requirement for extended prose are all very similar over the time reviewed. The reviewers found no change in how demanding the questions were from 2003 to 2008, but they thought that, in chemistry terms, the questions were more demanding when compared with those of other awarding organisations' in 2008.

WJEC's question papers were similarly demanding over the two years reviewed, apart from those in the foundation tier in 2003. The 2003 examination paper included questions which were more demanding for students and those in 2008, requiring them to demonstrate strategy skills and extended prose.

Edexcel's foundation-tier question papers were similarly demanding over the years reviewed, but the 2008 higher-tier question papers were less demanding than those in 2003 due to the questions being less complex. The requirement for numerical calculations increased at foundation level in 2008, although there was a slight decrease in the use of extended prose. The 2008 higher-tier question papers were also significantly less demanding because of a drop in questions requiring extended-prose answers.

AQA's foundation- and higher-tier question papers were similarly demanding over the time period reviewed. However, the nature of the tasks included in the 2008 scheme of assessment was more varied in terms of breadth, primarily because of the inclusion of the demands and challenges of How Science Works. There was some concern about the comparability of demand between the multiple-choice papers, Chemistry 1a and 1b, and the alternative structured paper, Chemistry 1.

The mark schemes from all of the awarding organisations, showed similar requirements and expectations when similar questions were compared. For example, in the 2003 higher-tier AQA paper C3 question 15b(i) is a bond-energy calculation and a very similar question 7b(i) in 2008 C3 paper (both are given 3 marks and similar answers are expected).

Given these variations, the demand within and across papers may differ depending on decisions made by the awarding organisation and, to a certain extent, by the

student in terms of their choice of assessment route (for example, the Edexcel scheme). When comparing 2003 with 2008, the reviewers thought that, overall, the AQA papers were more demanding, CCEA's were about the same, Edexcel's and WJEC's were slightly less demanding, and OCR's were the least demanding.

All papers are clear and well laid out either on a single page or double-page spread. There is suitable visual stimulus material throughout the majority of papers, although CCEA papers did not always match the other awarding organisations' in this area, and the effect was most marked for foundation-tier students. There was some concern that having both foundation- and higher-tier sections within the same examination paper may have confused some students (the controlled-response papers from Edexcel and AQA in 2008). However, overall it should be very clear to the students in both years what is required of them and how they are supposed to present their answers.

Coverage of objectives

The increase in the use of controlled-response and multiple-choice questions by Edexcel, OCR (and, to a certain extent, by AQA depending on the option taken), and the decrease in opportunity to use extended prose meant that there was very little scope in 2008 for students to demonstrate certain higher-order skills. For example, skills such as organisation of information, analysis, development of an argument (AO2a) and evaluation (AO2d) were not assessed sufficiently. These issues restricted the coverage of AOs that was possible in papers containing these types of question. However, taking into account the components and opportunities available in the scheme of assessment for internally assessed activities?

The assessment of How Science Works in both OCR and AQA was considered to be comprehensive.

In the AQA examination papers for both years and in both tiers there was greater coverage of AO2 compared with AO1. This is in marked contrast with WJEC and CCEA. However, there were some aspects of the AOs that were judged not to be addressed and matched fully in the operational papers in 2008 for AQA, WJEC and Edexcel, for example “understanding how decisions about science are made” and “evaluation of the impact of scientific developments” in (AO2c). The 2008 CCEA papers, albeit using the 2003 AOs, indicated a fairly narrow interpretation of some of the AOs, particularly “power and limitations of scientific ideas...” (AO1), and “apply principles to unfamiliar situations related to domestic, industrial and environmental contexts” (AO2).

Accessibility

The fact that a terminal-paper model has been replaced by a staged scheme of assessment, with shorter papers testing smaller amounts of content, has helped

improve accessibility for less able students. The clarity, language, layout and the visual-stimulus material was similar between the two years reviewed for all the awarding organisations except OCR, which was judged to have made the format more accessible, helping less able students. However, this has been slightly offset by the increased weighting for AO2, which has increased the potential demand for this group of students.

There was, in general, a decrease in the percentage of the marks allocated to numerical calculations in foundation-tier papers from 2003 to 2008. The percentage of marks allocated to writing symbol equations has not changed significantly and remains at a low level, although CCEA has a higher percentage. The percentage of marks allocated to extended prose has decreased except in the CCEA papers. The increase in reading time required in the AQA and Unit 3 OCR papers was judged to have made them more demanding for less able students. The one area of concern was Edexcel's 2008 C3 paper, where there is only one paper of 60 marks to cover the whole ability range. This meant a smaller percentage of marks accessible to less able students thus making the paper more demanding.

AQA's, Edexcel's and WJEC's papers appear to have been similarly demanding between 2003 and 2008, whilst OCR used an easier format and made their papers less demanding. CCEA's papers maintained a high level of demand for less able students.

Opportunities for high-attaining students

There were considerably fewer opportunities for students to use extended prose in their answers apart from in the AQA papers, when comparing 2003 with 2008. The Edexcel, WJEC and OCR papers in 2008 were judged to have very limited opportunities in this aspect. The opportunities for students to demonstrate their numerical skills, in terms of the percentage of marks allocated in the externally set papers, were slightly lower compared with 2003.

The high percentage of controlled-response questions and their structure in the 2008 OCR scheme of assessment (Unit 1 and 2 papers) has made the papers considerably less demanding in terms of the complexity of the questions and the use of the strategy skills required by students to answer them.

The Edexcel 2003 examination papers included an appropriate amount of material to challenge A and A* students. In the 2008 examination papers, however, there is less of this material, and the particular area of concern is paper C3, where there is only one paper of 60 marks. In addition, this paper has a weighting of 30 per cent to cover the whole ability range, which means a smaller percentage of marks available to discriminate and reliably grade A and A* students.

Those questions targeted at grade-A students on the higher tier in both years for AQA were very similar, for example those relating to empirical formula calculations and bond-energy determinations. The weighting of AO2 increased in the 2008 specification and there were increased challenges for higher-tier students, particularly on the C3 paper.

Those questions directed at grade-A students in the WJEC examination papers were very similar between 2003 and 2008 in terms of assessing similar topics. They required students to recall factual information and show understanding of scientific principles and concepts. However, although the weighting of AO2 had increased in 2008 reviewers were not convinced that there were more questions which required more application than recall or understanding when compared with the smaller weighting in 2003.

Coursework

Coursework requirements are outlined in [Appendix C](#). In 2003 (and in 2008 for CCEA), awarding organisations used a common framework for the assessment of practical skills. The Sc1 model was developed to assess practical and investigative skills using a common set of AOs and criteria. Sc1 was the only coursework component. In 2003 the Sc1 investigation was set internally, marked internally and moderated externally. The range of investigations used for the purposes of assessment became smaller over the years and restricted the opportunities for genuine investigative work. As centres gained more experience of the nature of Sc1, students achieved increasingly higher levels of attainment. This had the potential to decrease its discriminatory power and overcompensate for low examination marks, particularly for less able students.

The system in 2008 is much more complex and wide ranging with different assessment instruments and with different AOs adopted by awarding organisations. The skills and abilities that are internally assessed are related to practical work, although not necessarily investigative work, and to aspects of How Science Works.

The assessment models divide into two basic types with some compulsory and others optional, depending on the awarding organisation:

- externally set activities which are assessed internally
- internally set and assessed activities.

For AQA, the coursework comprises an investigative skills assignment. This assignment includes a practical activity followed by an externally set, internally assessed test related to the practical activity and taken in controlled conditions, and a practical skills assessment, which is holistically assessed over the course. The

practical skills assessment is not externally moderated, which is of concern in terms of ensuring validity and could be a non-discriminating part of the assessment.

Each investigative skills assignment has a novel context in the form of an externally set paper assessing the procedural content of How Science Works. Comparability between the different investigative skills assignments is addressed by using similar targeted or phrased questions. The papers are very demanding, particularly for the less able students. The reviewers thought that the investigative skills assignment and practical skills assessment combined together are of less demanding than the corresponding 2003 model.

For WJEC and OCR, there is a choice of the type of task to submit for assessment. WJEC students can offer a practical task, an extended report and/or an investigatory planning exercise. OCR students can submit either a case study and data-analysis exercise or an investigation. Comparability between these alternative routes is problematical and further scrutiny of students' work is necessary before any firm judgement can be made.

For Edexcel, internal assessment can vary from 30 per cent to 70 per cent of the marks for the whole scheme of assessment. Therefore, students can have a variable assessment experience and be exposed to variable demands. Some students will submit marks from centre-devised activities which are of a practical nature, whereas others will take written papers which are judged to assess different skills and abilities. Practical skills are assessed, as with AQA, using a holistic assessment over the course with the submitted marks not being subject to external moderation. The reviewers noted that 30 per cent of students achieved 100 per cent, and overall 70 per cent achieved an A*, suggesting that this component is not providing secure discrimination between students.

The increased range and complexity of the activities and tasks used in the internally assessed components makes it difficult to establish and judge clearly the effect on the level of demand on students. However, these components have the potential to be both accessible to weaker students and challenging for the most able.

Section 3: Standards of performance

Overview

Reviewers considered student work from all the awarding organisations from 2003 and 2008, and they made qualitative comments on the work they saw, predominately comparing the 2008 materials with each other. Details of the materials used can be found in [Appendix E](#), but not all the materials from the time period reviewed were available for use.

Findings

At the grade-A boundary, reviewers commented that for CCEA papers in 2008, a large breadth of the specification was covered, with many elements of chemistry tested. This breadth made it easier to ascertain how well students were able to apply concepts and processes. Overall reviewers found the paper to be more demanding than those of other awarding organisations. And they thought that the performance of the “weakest” grade-A CCEA student was greater than that of the best grade-A student of other awarding organisations.

For example, for Edexcel in 2008, the depth of challenge was thought not to probe sufficiently or discriminate between the most able students, and there were insufficient opportunities to apply abstract ideas over the breadth of the specification. Some opportunities were present to tackle calculations, but these were limited. There were also insufficient opportunities for students to explain abstract ideas using detailed terminology. The reviewers thought that these students did not have the same (more robust) knowledge and understanding as was present in the CCEA grade-A students’ scripts.

At the grade-C boundary on higher tier, reviewers commented that for AQA in 2008, students demonstrated more of the attributes required for the grade through their practical assessment than through the examinations. In the papers there were some questions that simply required common sense, rather than the ability to apply knowledge and understanding of chemistry at GCSE level. There was concern therefore, about the level of progression required from Key Stage 3.

At the grade-C boundary on foundation tier reviewers commented that AQA papers in 2008 were significantly better than Edexcel papers, with good breadth of coverage and response from students. Students were able to show skills, interpreting graphs and carrying out simple operations involving formulas. Furthermore, AQA students demonstrated the ability to write prose across a range of topics.

For CCEA in 2008, students demonstrated good facility in written prose, formula and some calculations, and a good descriptive knowledge of electrolysis. Students

demonstrated higher levels of knowledge and needed to have a greater knowledge of terminology than Edexcel students, and CCEA's students were thought to be better at explaining ideas. Edexcel students were not able to tackle more demanding aspects of the Edexcel papers, meaning that they could not demonstrate some of the characteristics expected at this level.

When comparing across the tiers, while there was some overlap between tiers, the foundation-tier students performed better than the higher-tier students. The students could show their skills better on foundation tier, for example their ability to write prose, use formulae and do simple numerical work. The higher-tier student work appeared to have many questions unanswered and negligible calculation work.

At the grade-F boundary student performance was consistent between the awarding organisations from the sample of scripts available.

Conclusions

This report details our analysis of the demand of qualifications across different years within GCSE chemistry. From the analysis, reviewers noted a number of elements that could usefully be considered when reviewing subject criteria within the subject, most particularly about the impact of altering the balance of AOs and the consequences of allowing variations in schemes of assessment. These findings fed into the review of subject criteria for the sciences at GCSE level and new criteria have now been introduced (published in 2009 for first teaching in September 2011).

The aims and learning outcomes in the new criteria are written specifically for chemistry, rather than being generic across all GCSE science subjects, and the AOs have changed significantly. These are detailed in the table on the following page.

Year introduced	Assessment objective		Weighting
2005	AO1	Knowledge and understanding of science and how science works	20–40%
2009	AO1	Recall, select and communicate their knowledge and understanding of chemistry	30–40%
2005	AO2	Application of skills, knowledge and understanding.	30–55%
2009	AO2	Apply skills, knowledge and understanding of chemistry in practical and other contexts	30–40%
2005	AO3	Practical enquiry and data-handling skills	20–40%
2009	AO3	And evaluate evidence, make reasoned judgements and draw conclusions based on evidence.	25–35%

Requirements for the proportion of external assessment have also been changed. A weighting of 75 per cent must now be allocated to external assessment and a weighting of 25 per cent to controlled assessment, whereas in 2008 a minimum of 25 per cent external assessment and a minimum of 25 per cent internal assessment weighting applied.

There is now a mathematics requirement within the chemistry subject criteria which did not appear in the previous science subject criteria. Candidates are now required to demonstrate a greater degree of mathematical knowledge, understanding and skills than was typically used in past GCSE science examinations.

Specifications developed on the basis of the new criteria will be subject to a further standards review in 2013/2014.

Appendix A: Provision of assessment materials and student work at GCSE and GCE levels for Ofqual's archive (annual inclusion and standards reviews)

Section 1: Specification of requirements

1.1 Each awarding organisation should draw the materials for each subject from the specification with their largest entry in summer 2008, unless that selection severely limits the range of examination components available. Where there are several entry options, materials should be drawn from the largest option only, unless Ofqual were exceptionally to agree other arrangements.

1.2 (With regards to GCSE) – where there are both modular and linear (non-modular) examinations in a subject, the awarding organisation operating the modular scheme with the greatest number of students (amongst all awarding organisations) should include that modular scheme, even if it is not a specification within the awarding organisation's largest entry. Similarly, the awarding organisation operating the linear scheme with the greatest number of students should include that linear scheme. If an awarding organisation runs both the largest entry linear examination and the largest entry modular examination in a subject, it will therefore provide two sets of materials, including student work, where required.

1.3 The following materials should be supplied:

a) Current specification: all associated question papers and final mark schemes.

b) The 2008 chief examiners' report (CER) and details of awarding procedures particular to the specification supplied.

c) An indication of how the specification's content and assessment criteria and objectives have been met in each question paper supplied. This may take the form of a grid. For objective tests this should include faculty values, discrimination indices and a specification grid detailing what grade each question was targeted at, as well as an indication of what percentage of students got a particular question correct when it was targeted at the grade they got overall.

d) Unit or component mark distributions (with grade boundary marks shown). It should be clear whether the marks are on the raw or uniform mark scale.

e) Grade boundaries, overall and by unit (both raw and scaled).

f) Student work as specified in Section 2.

g) Complete data record showing for each student selected the raw mark; final mark; weighted or uniform mark; grade for each component/unit (including any non-archived component/unit) and overall grade; and, where relevant, tier of entry.

Where appropriate, materials a)–e) may be supplied in electronic form.

Section 2: Student work

2.1 The work submitted should include the examination scripts, the internal assessment, and any oral/ aural examinations (with examiner mark sheet) where these are routinely recorded. In addition, for modular specifications, the examination papers of module tests should be supplied.

2.2 The sample should be of the original work of the students. Photocopies of work should only be used where it is impossible to send the originals and with agreement in advance by Ofqual. Student and centre names and numbers should be removed wherever they appear in a student's work, unless they form an integral part of the work, for example, within a letter.

2.3 Where an awarding organisation's specification has a relatively small entry or where, for some other reason, it is proving difficult to find sufficient students who fulfil the criteria, the awarding organisation should contact the Ofqual officer responsible to agree how best to finalise the sample.

2.4 All internal assessment submitted should be that of the particular students selected for the sample. If, for any reason, this proves to be impossible, the awarding organisation should contact the Ofqual officer responsible to agree appropriate alternative measures.

2.5 The sample of scripts retained for each specification (option) should be taken from students whose final mark lay at or near the subject grade boundaries for A/B, C/D and F/G for GCSE and A/B and E/U for GCE A level qualifications. At each boundary, each awarding organisation will supply the externally and internally set and marked assessments of fifteen students. Students selected should be those whose performance across units is not obviously and significantly unbalanced.

2.6 In tiered subjects, where the same grade boundary may feature in two tiers, separate sets of student work for the boundary should be provided from each tier.

In addition for AS/A level specifications:

2.7 Where awarding organisations have to supply student work for an A level specification, two samples are required: one for the AS and one for the A2 units.

2.8 For AS level, the work of 15 students whose mark for the AS is at or close to the UMS boundary for an AS grade A (240) or grade E (120) should be supplied.

Students selected should be those whose performance across the three AS units is not obviously or significantly unbalanced. Students should have taken at least two of the three AS units in the June examination series.

2.9 For A level, the sample comprises the A2 work of 15 students who have gained c240 UMS marks at A or c120 UMS marks at E on their A2 units. Students selected should be those whose performance across the three A2 units is not obviously or significantly unbalanced. Students selected will ideally have also gained an overall A level mark which is at or close to the UMS boundary for an overall A level grade A (480) or grade E (240). Students should have taken at least two of the three A2 units in the June examination series.

2.10 The set of AS and A2 units provided should also be a valid combination for A level.

2.11 Where coursework forms a compulsory sub-component within a unit, that coursework should also be collected. Where a unit has optional sub-components, the highest entry option should be supplied. The students chosen for the sample should, as far as possible, have a performance across the components of the unit which is not obviously unbalanced.

Appendix B: Assessment objectives

Year introduced	Assessment objective		Weighting
2000	AO1	Knowledge and understanding	45-55%*
2005	AO1	Knowledge and understanding of science and how science works	20–40%
2000	AO2	Application of knowledge and understanding, analysis and evaluation	25-35%
2005	AO2	Application of skills, knowledge and understanding	30–55%
2000	AO3	Investigative skills	20-25%
2005	AO3	Practical enquiry and data-handling skills	20–40%

*About one-third of which must be allocated to recall

Awarding organisation	Assessment objective	Weighting	
		2003	2008
AQA	AO1	50%	36%
	AO2	30%	44%
	AO3	20%	20%
CCEA	AO1	45–55%	45–55%
	AO2	25–35%	25–35%
	AO3	25%	25%
Edexcel	AO1	45–50%	20–40%
	AO2	25–35%	30–55%
	AO3	20–25%	20–40%
OCR	AO1	50%	30%
	AO2	30%	41%
	AO3	20%	29%
WJEC	AO1	45–55%	32–37%
	AO2	25–35%	30–43%
	AO3	20%	24–37%

Appendix C: Schemes of assessment

Awarding organisation	Tier	Number of papers		Time allocated (mins)	
		2003	2008	2003	2008
AQA	Foundation	1	3–4	135	3 x 45 = 135 or 2 x 30 + 2 x 45 = 150
	Higher	1	3–4	135	3 x 45 = 135 or 2 x 30 + 2 x 45 = 150
CCEA	Foundation	2	2	60+90 = 150	60+90 = 150
	Higher	2	2	90+120 = 210	90+120 = 210
Edexcel	Foundation	2	3–5	90+60 = 150	2 x 20 +30+20+60 = 150 max. or 2 x 20 + 20 = 60 min.
	Higher	2	3–5	90+60 = 150	2 x 20 +30+20+60 = 150 max. or 2 x 20+20 = 60 min.
OCR	Foundation	2	3	90+45 = 135	2 x 40 + 60 = 140
	Higher	2	3	90+45 = 135	2 x 40 + 60 = 140
WJEC	Foundation	1	3	120	3 x 45 = 135
	Higher	1	3	150	3 x 45 = 135

Assessment weightings of the extension material

Awarding organisation	2003		2008	
	Weighting (%)	Nature of assessment	Weighting (%)	Nature of assessment
AQA	25	no separate paper	25	45-minute written paper (foundation and higher papers)
CCEA	No change			
Edexcel	40	1-hour written paper	30	60-minute written paper (only one paper) or centre-devised activity
OCR	26.7	45-minute written paper	33.3	60-minute written paper (foundation and higher papers)
WJEC	25	no separate paper	25	45-minute written paper (foundation and higher papers)

Coursework: Internally assessed units for 2008

Awarding organisation	Type of assessment	Aim of assessment (practical + How Science Works)	Overall weighting	AO1	AO2	AO3
Edexcel	Compulsory, externally set, Internally assessed activities x 2	To explore How Science Works in context of units. One from Science + one from Additional. Science	10%	2–4%	8–11%	16–21%
	Compulsory internally set and assessed using broad published criteria Not moderated	To assess practical skills (follow instructions, collecting data, presenting results)	10%			
	Optional internally set and assessed External moderation	Quality of written communication, analysis of data, applications and implications of science	10%	1–2%	4–5%	4–5%
	Optional internally set and assessed External moderation	How Science Works, application of knowledge, practical enquiry	30%	5–16%	6–21%	0–11%

AQA	Compulsory Investigative skills assignment Externally set Internally assessed	Questions relate to analysis and evaluation of data + some aspects of How Science Works.	21.25%			
	Compulsory Practical skills assessment internally set and assessed using broad published criteria Not moderated	Holistic assessment of practical skills	3.75%		5%	20%
WJEC	Compulsory Three pieces of work (a) 3 x practical test (PT) externally set Internally assessed or (b) 2 x PT + extended report internally set and assessed external moderation or (c) 1 x PT + ER + Investigatory planning exercise internally set and assessed	Carrying out procedures, planning, analysis and evaluation. Alternative enquiry- based learning strategies, involving scientific and technological controversy Paper exercise on planning an experimental investigation	25%		5– 13%	12– 20%

	external moderation					
--	---------------------	--	--	--	--	--

OCR	Compulsory internally set and assessed using published criteria Case study (20%) and data analysis (13.3%) or a practical investigation (33.3%) External moderation	Study of a controversial issue in science (How Science Works + analysis and evaluation of data from practical work or a practical investigation	33.3%	2%	6%	25%
CCEA	Compulsory internally set and internally assessed using published criteria external moderation	An Sc1 practical investigation	25%			25%

Appendix D: GCSE specifications reviewed

GCSE 2003 and 2008					
Awarding organisation and specification codes					
Year	AQA	CCEA	Edexcel	OCR	WJEC
2003	3423	G14	1530	1981	012501-2
2008	4421	G14	2107	J634	023301

Appendix E: GCSE scripts reviewed

		Awarding organisation									
		AQA		CCEA		Edexcel		OCR		WJEC	
Year	Grade	2003	2008	2003	2008	2003	2008	2003	2008	2003	2008
		GCSE	A	8	8	8	8	8	8	8	8
C higher	8		4	4	8	8	4	8	4	8	4
C foundation	N/A		4	3	4	8	3	8	3	8	4
F	8		4	1	1	2	2	N/A	N/A	4	N/A

Appendix F: Availability of specification materials for the purposes of this review

Materials	2003 Materials					2008 Materials				
	GCSE					GCSE				
	AQA	CCEA	Edexcel	OCR	WJEC	AQA	CCEA	Edexcel	OCR	WJEC
Specification	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Question paper	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mark scheme	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Chief examiner's report	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mark distribution	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Grade boundaries	✗	✗	✗	✗	✗	✓	✓	✓	✓	✗
Assessment grids	✗	✗	✗	✗	✓	✓	✗	✗	✓	✗

✓ Material was available and was used in the review

✗ Material was not available and was not used in the review

Appendix G: Student achievement by grade

Awarding organisation and year	A*	A	B	C	D	E	F	G	U	Total student entries
AQA 2003	10%	17%	25%	28%	14%	3%	1%	0%	1%	5,902
AQA 2008	26%	29%	25%	15%	4%	1%	0%	0%	0%	43,217
CCEA 2003	0%	16%	33%	31%	16%	4%	1%	0%	0%	1,859
CCEA 2008	15%	32%	27%	19%	4%	2%	1%	0%	1%	2,647
Edexcel 2003	1%	16%	29%	26%	19%	6%	2%	0%	1%	5,923
Edexcel 2008	20%	28%	27%	17%	7%	1%	0%	0%	0%	9,298
OCR 2003	0%	29%	22%	18%	21%	9%	2%	0%	0%	9,263
OCR 2008	18%	33%	30%	14%	4%	1%	1%	0%	0%	10,743
WJEC 2003	0%	16%	24%	22%	26%	8%	2%	0%	1%	2,365
WJEC 2008	20%	28%	26%	21%	5%	15	0%	0%	0%	2,750

Appendix H: Review team

Review team		Organisation
Lead Reviewer	Geoff Mines	Ofqual reviewer
Specification Reviewers	Philip Eastwood	Ofqual reviewer
	Ian Hotchkiss	Ofqual reviewer
	Susan Walker	Ofqual reviewer
Script reviewers	Jane Bolton	Ofqual reviewer
	Stephen Plant	Ofqual reviewer
	John Donneky	AQA
	David Swann	Edexcel
	Dr David Lees	OCR
	Dr John Payne	Royal Society of Chemistry
	Sunetra Berry	The Association for Science Education

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