Ofqual

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Review of standards in summer 2009 GCSE Science and GCSE Additional Science

Joint Ofqual and DCELLS investigation

July 2010

Ofqual/10/4742

Contents

This report was finalised in March 2010 and shared with awarding organisations at that time. Specific references to awarding organisations submitting revised specifications to Ofqual in April 2010 should be read with that in mind.

Executive summary

In March 2009, Ofqual (the Office of Qualifications and Examinations Regulation) published a report detailing the findings of the monitoring of the first assessments of the new GCSE specifications in Science and Additional Science. The main findings were that standards of performance at the key grades in the sciences were lower than expected.

Since that report, Ofqual has carried out further monitoring work, and has worked with awarding organisations to review data used during the awarding process. In November 2009, we published data from the summer 2009 award, which showed that in GCSE Science both OCR A and Edexcel had significantly different results from those anticipated, suggesting that their standards were too low.

This report details the findings of a qualitative review carried out jointly by Ofqual and DCELLS of assessment materials and candidates' responses to the highest entry summer 2009 assessments from AQA, Edexcel, OCR and WJEC. Reviewers compared the work of candidates at key grade boundaries (A, C and F) with the published grade descriptions which set out the knowledge, understanding and skills that typical candidates with those grades are expected to demonstrate.

The findings largely support conclusions from earlier work. Candidate work from summer 2009 did not generally match the grade descriptions. Often this was due to the design of the question papers, which did not always give candidates of all abilities the opportunity to demonstrate their knowledge, understanding and skills.

These findings should be seen in the context of the current timetable for revised specifications in GCSE science subjects, which awarding organisations will submit for accreditation in April 2010. Ofqual, working with the regulators in Wales (DCELLS) and Northern Ireland (CCEA), will expect awarding organisations to take account of the recommendations in this report when developing their new specifications and assessment materials.

Introduction

Ofqual is the regulator of qualifications, examinations and assessments in England and vocational qualifications in Northern Ireland. Its work ensures that children, young people and adult learners get the results their work deserves, that standards are maintained and that the qualifications learners receive count now and in the future.

DCELLS (the Department for Children, Education, Lifelong Learning and Skills), a department of the Welsh Assembly Government, performs a similar role in Wales.

The awarding organisations that offer GCSE Science and Additional Science qualifications operate within a regulatory framework, which is set out in the following documents:

- *The statutory regulation of external qualifications in England, Wales and Northern Ireland* (QCA/04/1293; 2004) (www.ofqual.gov.uk/files/6944_regulatory_criteria_04(1).pdf)
- *GCSE, GCE, principal learning and project code of practice* (Ofqual/10/4718; 2010) (http://www.ofqual.gov.uk/public-download/category/53-codes-of-practice?download=680%3Acode-of-practice-2010).

The regulators carry out ad-hoc investigations into qualifications where concerns have been raised over the standards of assessments and/or the grades achieved by candidates. Such investigations normally include an analysis of the assessment materials and a review of the performance of candidates in a particular assessment.

An investigation of a particular qualification aims to:

- determine whether the assessments were fair and effective in measuring achievement by candidates in respect of the assessment objectives
- determine whether candidates were awarded grades for work that met the requirements of the published grade descriptions for the subject
- determine whether standards are comparable across awarding organisations
- provide advice to those developing revised criteria and/or specifications about what elements of the current assessments enable candidates to demonstrate what they know and how they can apply their knowledge

- identify any good practice that is worthy of encouragement and dissemination, to promote continuing improvement in the quality of examinations.

Each investigation involves a team of reviewers, who are contracted by the regulators and are experts in the subject being reviewed. These reviewers are sourced via consultant recruitment exercises carried out by the regulators, awarding organisation nominations, and individuals identified by subject associations and other learned organisations.

The team analyses the schemes of assessment, question papers, mark schemes and internally assessed tasks and review samples of candidate work, to determine whether it is in line with expectations set out in the published grade descriptions.

Background

New specifications for five separate GCSE science qualifications (Science, Additional Science, Biology, Chemistry and Physics) were introduced with first teaching in September 2006. This followed the revision of the national curriculum programme of study for science at key stage 4, in the light of widespread concern about the fitness for purpose of the previous science curriculum.

This revision was followed by the necessary accompanying revision of the GCSE science subject criteria1. Both were widely consulted upon during the development process. The science education and wider science communities, awarding organisations, learned societies, schools, colleges, universities and employers were among the wide range of stakeholders who contributed to the development of both the programme of study and the subject criteria.

The outcome was a new range of science GCSEs taught from September 2006, which aimed to ensure that:

- candidates were studying science that is relevant and up-to-date
- there was choice in the courses that students could take to prepare them for different career and/or learning routes after GCSEs.

The criteria for these new specifications provided scope for a range of different assessment approaches, including different weightings for external and internal

[1] GCSE applied science qualifications were also changed following this revision.

assessment, and placed a new emphasis on "How Science Works"2. The current schemes of assessment for Science and Additional Science 2009 specifications can be found in Appendix A.

Monitoring of new specifications in 2007 and 2008

A change in specifications can lead to difficulties in maintaining standards between the old and new qualifications. Ofqual monitored the new specifications in GCSE Science and GCSE Additional Science offered, by the three awarding organisations that are based in England (AQA, Edexcel and OCR) in 2007 and 2008.

In March 2009, Ofqual published the findings from these monitoring exercises: (www.ofqual.gov.uk/files/ofqual-09-4148_GCSE_science_2007_2008_report.pdf). The key findings were that:

- the wide variety of specifications, assessment routes and forms of assessment across awarding organisations led to candidates achieving the qualifications with very different sets of knowledge and skills being tested
- where specifications had a large number of discrete assessments with low weightings (for example, multiple objective tests) it was difficult to discriminate reliably between candidates of different abilities
- assessments did not always provide sufficient opportunity for more able candidates to demonstrate their scientific knowledge; in some questions, factors other than the demand of the scientific knowledge, understanding and skills were being assessed
- the assessment of the broader aspects of How Science Works was limited, with instances where the science being assessed did not depend on the candidates engaging with the context provided
- with regard to internal assessments, the questions on some sections of the practical assessments were not linked to the practical work and in some instances the marking criteria were unclear
- the sample of candidate work showed that the standard of work in the written papers was lower than expected and grade boundaries had been set at a relatively low proportion of marks available.

[2] How Science Works includes scientific method and the way scientific knowledge develops. It also focuses on the evidence (including the collection and interpretation of data) to support or refute ideas and theories that lie at the heart of science.

In light of the findings contained in the March 2009 report, the regulators asked awarding organisations to undertake the following:

- make improvements to the quality of questions, to provide appropriate opportunities for all candidates to show the extent of their abilities
- ensure that grade boundaries are in line with national standards
- implement tighter marking criteria to ensure that only the answers deserving of marks are credited
- undertake revisions of some internal assessment tasks to ensure better challenge to students and a closer link to the practical work
- where possible within the existing specifications, reduce the number of options available to candidates

When asking the awarding organisations to address these issues, it was acknowledged that bringing standards into line would take longer than a single year and summer 2009 would be the first step towards this goal. It was also acknowledged that, due to the production schedules for question papers taking up to 18 months from drafting a paper to candidates sitting it, it was unlikely that all of the issues relating to the quality of questions identified in the March 2009 report would have been addressed in the summer 2009 assessments.

Summer 2009 examinations

Ahead of the summer 2009 awards, Ofqual set out its expectations of the awarding organisations as to how to begin to address the decline in standards of performance at key grades identified as part of the monitoring of the 2007 and 2008 awards. In order that the summer 2009 examinations would be fair and that candidates would get the grade their work deserved, Ofqual referred the awarding committees to the recommendations in the March 2009 report. As units are awarded sequentially over the course of typically two years for a GCSE it was not possible for awarding organisations to implement all these changes in a way that would have an immediate effect on the qualification grades issued in GCSE science qualifications in summer 2009. However, the awarders were asked to:

- be circumspect and critical in their use of archive script evidence as candidates' work was not of an appropriate standard in 2007 and 2008 to guide awarding in 2009

- take into account that standards could be partly corrected by aligning the proportions of candidates achieving key grades in the summer 2009 awards with the most severe awards made in 2008.

In November 2009, Ofqual published data used to evaluate the outcome of the summer 2009 awards before the issue of results, together with data from the inter-awarding organisation screening carried out after the issue of results: (http://www.ofqual.gov.uk/files/2009-11-gcse-science-commentary.pdf). The data suggested that in GCSE Science, Edexcel and OCR had a significantly higher proportion of candidates achieving grades A and C, compared with other awarding organisations and predictions drawn from statistical indicators.

At the same time, Ofqual made clear that by summer 2010, we expect the differences between awarding organisations will have been substantially reduced by tightening standards further, where that is fair and appropriate.

This review

The data referred to above represent a statistical analysis of the summer 2009 results. Ofqual, at the same time, instigated an investigation into the qualitative evidence. GCSE Science and GCSE Additional Science are also taken by candidates in Wales. DCELLS was similarly concerned with standards in these qualifications, and this investigation was jointly commissioned by Ofqual and DCELLS. The scope of the 2009 review was extended to include WJEC, the awarding organisation based in Wales, as well as AQA, Edexcel and OCR. All four awarding organisations have candidates from England and Wales.

This review examined assessment materials and samples of work at the judgemental grade boundaries (those on which the awarding committees make their recommendations) in order to establish whether qualitative findings would support the statistical findings, and if so provide specific instances where these differences in demand were manifested. The findings of the review were also intended to help ensure that specifications and assessment models developed against the new subject criteria were more robust[3].

[3] Revised national criteria for GCSE Science and GCSE Additional Science were published in December 2009 (www.ofqual.org.uk/files/2009-12-04-gcse-subject-criteria-for-science.pdf and www.ofqual.org.uk/files/2009-12-04-gcse-subject-criteria-for-additional-science.pdf). Awarding organisations will be publishing new specifications for these qualifications and these will be used for teaching from September 2011.

Findings

The judgements made in this report reflect the outcomes of a two-stage process comprising an analysis of the assessment instrument (the question paper or practical task, mark scheme and associated documentation) and an assessment of candidate achievement primarily at unit level. These were undertaken using the specification with the largest number of entries in summer 2009 from each awarding organisation[4]. The study was conducted by subject-specialist reviewers contracted by Ofqual and DCELLS.

Assessment materials and candidate work at grades A, C and F were compared against the grade descriptions (see Appendix D) which describe the typical performance of a candidate at that grade at qualification level. This investigation was carried out in autumn 2009 using script evidence from the summer 2009 assessments. Full details of the methodology can be found at Appendices B and C.

The findings of this investigation did not differ significantly from those found in previous investigations, thus adding further evidence that standards are currently too low in GCSE Science and Additional Science qualifications. However, the performance of assessments varied widely across awarding organisations and within units for individual awarding organisations.

The findings for this study can be broadly grouped around the following issues:

- the level of demand of the assessments
- the comparability of assessments within specifications and between awarding organisations
- the effectiveness of coverage of the assessment objectives
- the consistent application of mark schemes.

In expanding the findings under these headings, the examples illustrate the evidence base used. These examples do not cover all instances of evidence from the identified or other awarding organisations or specifications.

[4] With the exception of AQA Science, where both specifications (A and B) were reviewed as they offered very different types of assessment.

Level of demand

For both Science and Additional Science, the level of demand of the assessments and the standard of candidate work was at an acceptable level for AQA and WJEC, while the Edexcel and OCR assessments were considered to have too low a level of demand. Across all awarding organisations some grade boundaries were too low to ensure that candidates showed a satisfactory range of knowledge and understanding across the assessed content. For example, in some Additional Science assessments, the boundary mark for grade A was sometimes less than half of available marks, and the boundary mark for grade C was as low as 6 out of 30 marks in one case[5].

A low level of demand in an assessment was characterised by insufficient opportunities for candidates to demonstrate how they could meet the requirements for a particular grade. There were various reasons for this: the nature of the questions did not allow for a sufficient range of content and skills to be covered; question demand was derived from elements in the questions other than science content; or there was insufficient challenge to enable the paper as a whole to discriminate between candidates of different abilities.

Externally assessed units

Where assessments were of an acceptable level of demand, they enabled candidates with a range of achievements to access questions and to demonstrate their knowledge, apply their knowledge and use their skills. For example, in GCSE Science, candidates at the A/B boundary for AQA and WJEC generally demonstrated good knowledge and understanding of science content and a good range of skills in the application and interpretation of data.

However, there were many instances of demand being at an inappropriate level. For example, in OCR Science A Unit 2 (A212), many of the questions could be answered from an English comprehension of the stimulus materials given in the question.

In some cases, there were insufficient questions to offer high-ability candidates the opportunity to demonstrate performance at a grade A standard. For example, the Edexcel Science biology assessment (5002) and OCR Science A Unit 1 (A211) were of relatively low demand and failed to provide sufficient challenge for the most able candidates. Specifically, in OCR Science A Unit 1, the chemistry questions were judged to be of low demand and did not provide the necessary stretch and challenge

[5] Low boundary marks in assessments can in some instances support the maintenance of standards at qualification level where there are a large number of assessments (the 'regression' effect). In these cases it was seen to compromise the amount of evidence candidates needed to demonstrate to achieve a particular grade.

for the highest-ability candidates. Generally, where assessments relied heavily on objective tests (particularly AQA Science A, Edexcel Science and OCR Science A and Additional Science A) there were insufficient opportunities for candidates to demonstrate written communication or for candidates to use technical vocabulary. Although technical vocabulary was used in the questions, there was no requirement for candidates to use it themselves.

In some cases, question papers did not address key parts of the assessment objectives, which limited the opportunities for candidates to demonstrate fully the requirements of the grade descriptions. For example, in the OCR Science A Unit 3 assessment, there was little or no opportunity to demonstrate awareness of specification content such as 'why decisions about uses of science are made in familiar contexts' across the papers and 'apply this to unfamiliar circumstances' therefore making it less challenging at grade A.

With the exception of AQA, the standard demonstrated at grade A at subject level for Additional Science was too low. Candidates failed to show the required depth of scientific knowledge and understanding and did not have the opportunity to demonstrate higher skills such as writing in extended prose or carrying out meaningful calculations.

Internally assessed units

Providing adequate stretch in assessments was also an issue in the internally assessed units. For example, in GCSE Science, the AQA ‘Investigative Skills Assignment’ and the WJEC ‘Practical task’ were judged to provide appropriate opportunities to assess the broad range of skills required, although the structured nature of the questions limited the opportunity for candidates to display some of the higher-level skills.

In Additional Science there were issues with demand across awarding organisations. The OCR Additional Science A 'Practical' was more demanding than those of other awarding organisations due to the open ended nature of the task.

Relative demand of foundation and higher tier units

There were also concerns about the comparability of demand between tiers within awarding organisations. These were most notable in assessments offered by AQA and Edexcel in both Science and Additional Science.

For example, while questions targeted at the higher tier candidates across AQA Science A assessments provided opportunity for candidates to demonstrate knowledge of concepts, techniques and facts across most areas of the specification, the questions in the foundation tier biology, targeting grade C, were of a much lower

demand than those on the higher tier questions. Clearly the overall level of demand of a foundation tier paper will be lower than that of a higher tier paper. However, candidates achieving a grade C on the foundation tier questions demonstrated a lower level of achievement (in relation to the grade descriptions for grade C) than candidates achieving a grade C on the higher tier set of questions. Similar issues were identified in the AQA Additional Science papers, with a difference seen in the level of demand of the questions and in candidate performance across the tiers to gain a grade C. Candidates achieving a grade C on the foundation tier units did not demonstrate the same level of knowledge and understanding as candidates achieving a C on the higher tier units.

Another example of this was found in the Edexcel Science biology paper 1b where a high proportion of the foundation tier questions tested knowledge and understanding of the specification at a relatively low level, making it relatively undemanding for grade C candidates. In the Edexcel Additional Science papers the difference in style between foundation and higher tier papers was significant. Questions on the foundation tier tended to be more closed and rely more on multiple choice than in the higher paper where questions were much more open and challenging.

Coverage of assessment objectives

There were a number of elements of the assessment objectives that were consistently under-represented for both GCSE Science and Additional Science in the assessments. These were:

- planning
- interpreting data/numerical skills
- assessment of How Science Works
- the use of technical language and communication.

Planning

Across all awarding organisations' assessments, opportunities for candidates to demonstrate planning skills were limited.

For example, across AQA Science A, generally, there were few opportunities for candidates to plan a scientific task, although biology B1b involved aspects of planning and understanding of the procedures involve in planning. This was also the case across AQA Science B and for OCR Science A Unit 1 (A211). Similarly, in the Edexcel Additional Science assessments there was no opportunity to plan a scientific task, although candidates were required in the physics papers to establish a method and plan what to do.

In the internally assessed units for both Science and Additional Science for AQA, Edexcel and WJEC and the internally assessed unit for OCR Science A, planning skills were not adequately assessed. As a result, candidates could not demonstrate planning as required in the grade descriptions for grades A, C and F.

However, planning was evident in some assessments. For example, OCR Science A Unit 3 included a planning task in the physics questions and OCR Science A Unit 4 required candidates to plan a scientific task. The chemistry higher tier paper in WJEC Science also included a planning task.

Further, the OCR Additional Science A 'Practical' required candidates to carry out a single non-prescribed, open-ended practical investigation of the candidate's choice. This included an element of planning that was stronger than other internal assessments.

Interpreting data/numerical skills

There were relatively few opportunities to demonstrate numerical skills across all the specifications, and the demand, particularly for the most able candidates, was limited. However, there were some examples of good practice.

In OCR Science A unit 4 and WJEC Science there were opportunities for candidates to interpret data from secondary sources given in the questions and identify patterns.

Across AQA Science A units there were opportunities for candidates to interpret data from secondary sources given in the questions, present those data in appropriate ways, identify patterns and draw conclusions. AQA Science A papers provided candidates with some opportunities to demonstrate numerical skills. However, since this was an objective test, candidates could not demonstrate numerical fluency. Across AQA Science B units there were opportunities for candidates to demonstrate numerical skills in all but the chemistry higher tier papers. In general, there was more evidence of appropriate levels of demand of numerical skills in assessments provided by AQA and WJEC.

How Science Works

The assessment of How Science Works varied between awarding organisations. While OCR provided good coverage, some awarding organisations, notably WJEC, placed less emphasis on the assessment of How Science Works, especially in the physics and chemistry papers.

AQA assessments across Science and Additional Science had limited assessment of How Science Works. The biology papers for AQA Science B (BLY1F and BLY1H) provided the best evidence of achievement in How Science Works of all AQA papers.

Technical vocabulary and communication

Candidates' use of technical vocabulary and communication skills generally matched the grade descriptions across the AQA Science B and WJEC written papers in Science, but were not adequately tested on the OCR Science A and OCR Additional Science A (with the exception of unit 4), AQA Science A and Edexcel Science external assessments, due to the emphasis on objective test questions.

These assessments provided no opportunities for candidates to demonstrate communication skills and technical vocabulary and techniques. Although technical vocabulary may have been used in the questions there was no requirement for candidates to use technical vocabulary independently. In addition, language and fluency could not be tested and communication skills were limited to reading and understanding the questions.

Inconsistent marking

Inconsistency of marking was of particular note in the internally assessed units for Additional Science. For example, while OCR's model for the internally assessed unit in Additional Science (A220) enabled candidates to plan an experiment, the open-ended nature of the task, the lack of prescription and the use of a generic mark scheme led, in some cases, especially at the upper end of the achievement scale, to variability in marking. Assessors struggled to interpret marking scheme statements, in particular those which required assessors to make judgements on the complexity and demand of task, the techniques used and the autonomy displayed by candidates.

There was also evidence of inconsistent marking in the internally assessed WJEC Additional Science unit, which sometimes led to the award of similar grades for a range of dissimilar quality answers.

Recommendations for action

Awarding organisations are currently developing specifications and sample assessment materials to meet the revised subject criteria for GCSE sciences (http://www.ofqual.gov.uk/public-download/category/53-codes-of-practice?download=680%3Acode-of-practice-2010and www.ofqual.gov.uk/files/2009-12-04-gcse-subject-criteria-for-additional-science.pdf).

Awarding organisations should take account of the following recommendations for current specifications as well as for new specifications being developed.

1. Question papers in the new specifications should be constructed to provide more appropriate balance between objective test questions, short answer questions and open response questions, in order to allow candidates of all abilities to demonstrate their knowledge, understanding and skills, particularly at higher levels

2. Question papers and internally assessed tasks in the new specifications, across all routes, should allow all candidates to demonstrate their written communication skills, including their use of technical vocabulary

3. All questions in the assessments, across all routes, should assess *scientific* knowledge, understanding and skills and assessments should ensure appropriate coverage of How Science Works across the specification

4. Question papers and internal assessments should cover all aspects of targeted assessment objectives to enable candidates to demonstrate all aspects of the grade descriptions

5. Assessment setters should make use of feedback from previous examination series to ensure that papers are constructed so that grade boundaries are not set at inappropriately low marks compared to the demand of the assessment

6. Question papers must ensure comparability of performance to achieve a grade C irrespective of tier.

Next steps

Ofqual, working with the regulators in Wales and Northern Ireland, will continue its monitoring work in relation to GCSE science. We will attend many of the grade awarding meetings in 2010 and we will facilitate discussions with awarding organisation technical staff about how best to evaluate the results.

To ensure that actions from this and other reports are met in the new specifications being developed for first teaching in 2011, the regulators will closely scrutinise sample assessment materials submitted in the accreditation of the new specifications to ensure that:

- they represent a valid assessment model for How Science Works, and
- individual assessments have sufficient scope to discriminate accurately across the full range of candidate performance.

The regulators will also require the awarding organisations to ensure that new specifications attend to the views expressed by the science community about increasing the mathematical demands in new GCSE examinations in science subjects.

When scrutinising sample assessment materials and, subsequently, live assessments, the regulators will need to be satisfied that awarding organisations require candidates to demonstrate a greater degree of mathematical knowledge, understanding and skills than is typically used in current assessments for GCSE science qualifications.

Appendix A: Schemes of assessment

AQA GCSE Science A (4461)

Summary: Externally assessed units – six objective test papers, with separate papers for chemistry, physics and biology. Internally assessed element – a classroom practical followed by an externally set, internally assessed test and a practical skills assessment.

Unit title (code)	Weighting %	Time allowance	Marks	Type of assessment
Biology 1a (BLY1AP)	12.5	30 minutes	36	Objective test – available as paper-based or on-screen test in centres. Foundation and Higher Tier questions in common paper.
Biology 1b (BLY1BP)	12.5	30 minutes	36	Objective test – available as paper-based or on-screen test in centres. Foundation and Higher Tier questions in common paper.
Chemistry 1a (CHY1AP)	12.5	30 minutes	36	Objective test – available as paper-based or on-screen test in centres. Foundation and Higher Tier questions in common paper.
Chemistry 1b (CHY1BP)	12.5	30 minutes	36	Objective test -available as paper-based or on-screen test in centres. Foundation and Higher questions in common paper.
Physics 1a (PHY1AP)	12.5	30 minutes	36	Objective test – available as paper-based or on-screen test in centres. Foundation and Higher Tier questions in common paper.
Physics 1b (PHY1BP)	12.5	30 minutes	36	Objective test – available as paper-based or on-screen test in centres. Foundation and Higher Tier questions in common paper.
Science Centre-Assessed Unit (SCYC)	25	45 minutes	40	Externally set, internally assessed and practical assessment

AQA GCSE Science B (4462)

Summary: Externally assessed units – three structured papers, with separate papers for chemistry, physics and biology. Internally assessed unit – a classroom practical followed by an externally set, internally assessed test and a practical skills assessment.

Unit title (code)	Weighting %	Time allowance	Marks	Type of assessment
Biology 1 (BLY1F or BLY1H)	25	45 minutes	45	Short answer written test. Foundation and Higher Tier tests available.
Chemistry 1 (CHY1F or CHY1H)	25	45 minutes	45	Short answer written test. Foundation and Higher Tier tests available.
Physics 1 (PHY1F or PHY1H)	25	45 minutes	45	Short answer written test. Foundation and Higher Tier tests available.
Science Centre-Assessed Unit (SCYC)	25	45 minutes	40	Externally set, internally assessed and practical assessment.

AQA GCSE Additional Science (4463)

Summary: Externally assessed units – three structured papers, with separate papers for chemistry, physics and biology. Internally assessed unit – a classroom practical followed by an externally set, internally assessed test and a practical skills assessment.

Unit title (code)	Weighting %	Time allowance	Marks	Type of assessment
Biology 2 (BLY2F or BLY2H)	25	45 minutes	45	Short answer written test. Foundation and Higher Tier tests available.
Chemistry 2 (CHY2F or CHY2H)	25	45 minutes	45	Short answer written test. Foundation and Higher Tier tests available.
Physics 2 (PHY2F or PHY2H)	25	45 minutes	45	Short answer written test. Foundation and Higher Tier tests available.
Additional Science Centre-Assessed Unit (ASCC)	25	45 minutes	40	Externally set, internally assessed and practical assessment.

Edexcel GCSE Science (2101)

Summary: Externally assessed units – six papers, with separate papers for chemistry, physics and biology. Internally assessed unit – three assessment activities (one for biology, one for chemistry and one for physics) and practical skills assessment.

Unit title (codes)	Weighting %	Time allowance	Type of assessment	Comments
Science (5001, 5002, 5003, 5004, 5005, 5006, 5007, 5008, 5009, 5010)	40% – Internal assessment 60% – External assessment	45 minutes – Internal assessment 20 minutes – External assessment	Compulsory internally assessed plus six externally assessed	Internal assessment (40%) • Assessment of practical skills (10%), where the teacher is assessing the student's ability to follow instructions, collect data (by taking readings and measurements, making observations and by using ICT) and present their raw results. Not moderated. • Assessment activities (3 x 10%), devised by Edexcel, marked by the teacher and externally moderated by an examiner appointed by Edexcel. External assessment (60%) • 60% based on six tiered multiple-choice tests available in November, March and June.

Edexcel Additional Science (2103)

Summary: Externally assessed units – nine papers with one objective question paper each for chemistry, physics and biology. The other six papers are structured papers, with separate papers for physics, chemistry and biology. Internally assessed unit – three assessment activities (one for biology, one for chemistry and one for physics) and practical skills assessment.

Unit title (codes)	Weighting %	Time allowance	Type of assessment	Comments
Additional Science (5011, 5012, 5013, 5014 plus two of: 5015, 5016 F or H, 5021 plus two of: 5017, 5018 F or H, 5022 plus two of: 5019, 5020 F or H, 5023)	40% - Internal assessment 60% - A choice of further assessment routes available	45 minutes – Internal assessment 20 minutes each – Externally assessed multiple-choice tiered test 30 minutes each – Externally assessed structured tiered examination paper	Compulsory internally assessed plus two externally assessed	Internal assessment (40%) • Assessment of practical skills (10%), where the teacher assesses student's ability to follow instructions, collect data (by taking readings and measurements, making observations and by using ICT) and to present their raw results. Not moderated. • Assessment activities (3 x 10%), devised by Edexcel, marked by the teacher and externally moderated by an examiner appointed by Edexcel. A choice of further assessment routes available (60%) Students must attempt two out of the three routes for each of biology, chemistry and physics; each assessment contributes 10%: • internally assessed centre-devised • externally assessed multiple-choice tiered tests available in November, March and June • externally assessed structured tiered examination papers available in November, March and June.

OCR GCSE Science A (J630)

Summary: Externally assessed units – four composite papers (three objective tests, one structured) made up of physics, chemistry and biology questions. Internally assessed unit – a practical data analysis (does not require practical work) and a case study.

Unit title (code)	Weighting %	Time allowance	Marks	Type of assessment
Unit 1 – modules B1, C1, P1 (A211)	16.7	40 minutes	42	Objective test
Unit 2 – modules B2, C2, P2 (A212)	16.7	40 minutes	42	Objective test
Unit 3 – modules B3, C3, P3 (A213)	16.7	40 minutes	42	Objective test
Unit 4 – Ideas in Context (A214)	16.7	45 minutes	40	Short answer written test
Unit 5 – Practical Data Analysis and Case Study (A219)	33.3 (Split between 13.3% and 20%)	Not allocated	40 (16+24)	Skills assessment

OCR GCSE Additional Science A (J631)

Summary: Externally assessed units – four composite papers (three objective tests, one structured) made up of physics, chemistry and biology questions. Internally assessed unit – single non-prescribed, open-ended practical investigation of the candidate's choice.

Unit title (code)	Weighting %	Time allowance	Marks	Type of assessment
Unit 1 – modules B4, C4, P4 (A215)	16.7	40 minutes	42	Objective test
Unit 2 – modules B5, C5, P5 (A216)	16.7	40 minutes	42	Objective test
Unit 3 – modules B6, C6, P6 (A217)	16.7	40 minutes	42	Objective test
Unit 4 – Ideas in Context (A218)	16.7	45 minutes	40	Short answer written test
Unit 5 – Practical (A220)	33.3	Not allocated	40	Skills assessment

WJEC GCSE Science (1310)

Summary: Externally assessed units – separate structured papers for biology, chemistry and physics. Internally assessed unit – practical skills assessment.

Unit title (no unit codes)	Weighting %	Time allowance	Marks	Type of assessment
Science – Biology 1	25	45 minutes	50	Short answer written test. Foundation and Higher Tier tests available.
Science – Chemistry 1	25	45 minutes	50	Short answer written test. Foundation and Higher Tier tests available.
Science – Physics 1	25	45 minutes	50	Short answer written test. Foundation and Higher Tier tests available.
Science – (B+C+P)	25	N/A	75	Practical work provided by WJEC or written by centres for completion wholly during class time and marked by teachers to a mark scheme provided/approved by WJEC or a centre assessed extended report.

WJEC Additional Science (1370)

Summary: Externally assessed units – separate structured papers for biology, chemistry and physics. Internally assessed unit – practical skills assessment.

Unit title (no unit codes)	Weighting %	Time allowance	Marks	Type of assessment
Additional Science – Biology 2 (239/01)	25	45 minutes	50	Short answer written test. Foundation and Higher Tier tests available.
Additional Science – Chemistry 2 (240/01)	25	45 minutes	50	Short answer written test. Foundation and Higher Tier tests available.
Additional Science – Physics 2 (241/01)	25	45 minutes	50	Short answer written test. Foundation and Higher Tier tests available.
Additional Science – (B+C+P)	25	N/A	75	Internally assessed

Appendix B: Methodology

Provision of assessment materials and candidate work

Each awarding organisation was asked to provide assessment materials for GCSE Science and GCSE Additional Science from the specification with their largest entry in summer 2009. The materials requested included: the current specification, all associated question papers, final mark schemes, the 2009 examiners' report and grade boundaries, overall and by unit (both raw and scaled).

The sample of scripts requested for each specified unit comprised candidate work where the final mark lay at or near the unit grade boundaries for A/B, C/D and F/G. Details of the Ofqual/DCELLS requirements for assessment materials and candidate work can be found in Appendix C.

The review team

Twenty reviewers, experts in at least one of biology, chemistry or physics, were contracted by Ofqual/DCELLS to undertake the review. These reviewers were sourced via a consultant recruitment exercise carried out by Ofqual for previous reviews, awarding organisation nominations, and individuals identified by subject associations and other learned organisations. There was a balance of reviewers associated with each of the awarding organisations, subject associations and those with no declared association.

Analysis of the assessment instrument

The reviewers were asked to familiarise themselves with the specifications and question papers for either Science or Additional Science. They were each allocated question papers from two awarding organisations in order to review these against the published GCSE grade descriptions for grades A, C and F. They were asked to comment on the following:

- questions that gave opportunities for responses to match the grade description
- questions where there was no opportunity for responses to match the grade description
- questions that were either too demanding for grade A or not demanding enough for grade F, or vice versa.

Assessment of candidate performance

Once familiar with the specifications, grade descriptions and question papers the reviewers were brought together for a two-day residential meeting in order to analyse candidate performance (a script review). This started with a briefing session to ensure that all the reviewers had a common understanding of the methodology and the judgement criteria, and to give the reviewers an opportunity to discuss observations from their initial review of assessment materials.

The scripts (candidate work) were organised into packs for consideration during the review. For the externally assessed units, each pack contained ten scripts at the same grade from a single awarding organisation in a single specialisation (biology, chemistry or physics). Each pack of internally assessed work contained scripts from a single awarding organisation at the same grade, but due to the nature of the internally set assignments, the packs were made up of work across specialisations (ie they included work covering some or all of biology, chemistry and physics). The number of scripts varied according to the material available.

The review was structured so that each reviewer saw all materials from all awarding organisations at all grades for their specialisation for the externally assessed materials, and all packs of internally assessed work across all specialisations.

For the packs containing externally assessed work reviewers were asked to rate the pack as a whole using a five-point scale to judge how well the candidate work met the relevant grade descriptions. They were also asked to record any comments they had about individual scripts.

For the packs containing internally assessed work reviewers were again asked to rate the pack as a whole using a five-point scale to judge how well the candidate work met the relevant grade descriptions. Additionally, they were asked to take into account how well the assessment was targeted at the identified assessment objectives.

The scale used to score the packs was:

1 = Appropriate; 2 = Acceptable; 3 = Limited; 4 = Poor; and 5 = Inappropriate

Additional review of Edexcel objective tests

All of the candidate work for Edexcel Science and half of the Additional Science comprised responses to objective test questions. Consequently, only the machine readable answer sheets were available for these parts of the assessments.

An additional home-working exercise was commissioned for reviewers to comment on the statistical analysis carried out by Edexcel on the outcome of their tests. This

allowed reviewers to comment on the overall outcome of the test rather than reviewing the responses of individual candidates.

The questions from the Science and the Additional Science examination papers were rearranged in facility order, starting with the question answered correctly by most candidates. The grade boundaries for A, C and F were indicated on this rearranged order and reviewers were asked to comment on the following:

- the questions that those just getting the grade got correct
- the questions that were answered incorrectly by those just getting the grade
- any questions that appeared to perform erratically.

The reviewers submitted their comments in report form and these were collated and summarised.

Limitations of the review

This review differed from an established standards review in that:

- it was a review of one awarding series and the objective was to measure the comparability of the grades awarded across the awarding organisations in that particular series
- the reviewers were asked to judge whether the scripts reviewed were worthy of the grade awarded, when compared to the grade description for that grade, rather than ranking them or comparing them against another script.

The review relied heavily on the judgments made by the reviewers, and human judgements cannot be completely objective and instructions given can be open to misinterpretation.

In an effort to mitigate the inherent unreliability in the process the following steps were taken:

- experienced subject reviewers were used with a balance between reviewers with affiliations with awarding organisations and those with no affiliations
- efforts were made to maximise the number of reviewers providing decisions, to increase the accuracy of the findings

- reviewers were thoroughly briefed to ensure a common understanding of the methodology and the judgement criteria
- documentation was constructed to assist the judgement process.

All the above steps were vital in this particular review. The judgements were mainly qualitative, and so the comments that reviewers made were very important when reporting the findings.

In the main, this review generated qualitative data, which limited the amount of statistical analysis that could be carried out. The only opportunity for statistical analysis was for a straightforward collation of the ranking scores awarded to each of the packs.

The statistical data gives a limited overview of the relative ranking of awarding organisations and the appropriateness of their grade boundaries against the published grade descriptions, but the qualitative judgements allow a more in-depth analysis of the performance.

Appendix C: Provision of assessment materials and candidate work

Section 1 Specification of requirements

1.1 Each awarding organisation should draw the materials for each subject from the syllabus with their largest entry in summer 2009, 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 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 candidates (amongst all awarding organisations) should include that modular scheme, even if it is not a syllabus within the awarding organisation's largest entry. Similarly, the awarding organisation operating the linear scheme with the greatest number of candidates 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 candidate work, where required.

1.3 The following materials should be supplied:

a) Current syllabus: all associated question papers and final mark schemes.
b) The 2009 examiners' report and details of awarding procedures particular to the syllabus supplied.
c) An indication of how the syllabus’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 candidates 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) Candidate work as specified in Section 2.
g) Complete data record showing for each candidate 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 Candidate 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 syllabuses, the examination papers of module tests should be supplied.

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

2.3 Where an awarding organisation's syllabus has a relatively small entry or where, for some other reason, it is proving difficult to find sufficient candidates 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 candidates 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 syllabus (option) should be taken from candidates whose final mark lay at or near the subject grade boundaries for A/B, C/D and F/G. At each boundary, each awarding organisation will supply the externally set and marked assessment of ten candidates; for five of these candidates, the internally set and marked assessment should also be included. The candidates 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 candidate work for the boundary should be provided from each tier.

Section 3 Collating candidates' work

3.1 Within each subject in each year, all the work of a candidate should be put together under cover of a standard sheet which includes the following information:

- 2009

- GCSE
- Subject (Science or Additional Science)
- Grade (If GCSE tiered examination, add letter in bracket to indicate tier of entry: ‘F’ for foundation, ‘I’ for intermediate, ‘H’ for higher)
- Candidate number (Use from 1 to 10 for each grade boundary in each subject).

Appendix D: Assessment objectives and grade descriptors for GCSE Science and Additional Science

Extracted from 'Criteria for GCSE Science':
www.qcda.gov.uk/libraryAssets/media/11881_gcse_science_criteria_apr05.pdf

5. ASSESSMENT OBJECTIVES

5.1 Specifications must require that all candidates demonstrate the following assessment objectives in the context of the prescribed skills, knowledge and understanding. Within the assessment objectives, candidates should be required to use communication skills, including ICT, as specified in 3.6 (iii), using scientific conventions (including chemical equations) and mathematical language (including formulae), where appropriate.

5.2 **Assessment objective 1 (AO1): Knowledge and understanding of science and how science works**

Candidates should be able to:

a) demonstrate knowledge and understanding of the scientific facts, concepts, techniques and terminology in the specification

b) show understanding of how scientific evidence is collected and its relationship with scientific explanations and theories

c) show understanding of how scientific knowledge and ideas change over time and how these changes are validated.

5.3 **Assessment objective 2 (AO2): Application of skills, knowledge and understanding**

Candidates should be able to:

a) apply concepts, develop arguments or draw conclusions related to familiar and unfamiliar situations

b) plan a scientific task, such as a practical procedure, testing an idea, answering a question, or solving a problem

c) show understanding of how decisions about science and technology are made to different situations, including contemporary situations and those raising ethical issues

d) evaluate the impact of scientific developments or processes on individuals, communities or the environment.

5.4 **Assessment objective 3 (AO3): Practical, enquiry and data-handling skills**

Candidates should be able to:

a) carry out practical tasks safely and skilfully

b) evaluate the methods they use when collecting first-hand and secondary data

c) analyse and interpret qualitative and quantitative data from different sources

d) consider the validity and reliability of data in presenting and justifying conclusions.

7. GRADE DESCRIPTIONS

Grade F

Candidates demonstrate a limited knowledge and understanding of science content and how science works. They use a limited range of the concepts, techniques and facts from the specification, and demonstrate basic communication and numerical skills, with some limited use of technical terms and techniques.
They show some awareness of how scientific information is collected and that science can explain many phenomena.
They use and apply their knowledge and understanding of simple principles and concepts in some specific contexts. With help they plan a scientific task, such as a practical procedure, testing an idea, answering a question, or solving a problem, using a limited range of information in an uncritical manner. They are aware that decisions have to be made about uses of science and technology and, in simple situations familiar to them, identify some of those responsible for the decisions. They describe some benefits and drawbacks of scientific developments with which they are familiar and issues related to these.
They follow simple instructions for carrying out a practical task and work safely as they do so.
Candidates identify simple patterns in data they gather from first-hand and secondary sources. They present evidence as simple tables, charts and graphs, and draw simple conclusions consistent with the evidence they have collected.

Grade C

Candidates demonstrate a good overall knowledge and understanding of science content and how science works, and of the concepts, techniques, and facts across most of the specification. They demonstrate knowledge of technical vocabulary and techniques, and use these appropriately. They demonstrate communication and numerical skills appropriate to most situations.
They demonstrate an awareness of how scientific evidence is collected and are aware that scientific knowledge and theories can be changed by new evidence.
Candidates use and apply scientific knowledge and understanding in some general situations. They use this knowledge, together with information from other sources, to help plan a scientific task, such as a practical procedure, testing an idea, answering a question, or solving a problem.
They describe how, and why, decisions about uses of science are made in some familiar contexts. They demonstrate good understanding of the benefits and risks of scientific advances, and identify ethical issues related to these.
They carry out practical tasks safely and competently, using equipment appropriately and making relevant observations, appropriate to the task. They use appropriate methods for collecting first-hand and secondary data, interpret the data appropriately, and undertake some evaluation of their methods.
Candidates present data in ways appropriate to the context. They draw conclusions consistent with the evidence they have collected and evaluate how strongly their evidence supports these conclusions.

Grade A

Candidates demonstrate a detailed knowledge and understanding of science content and how science works, encompassing the principal concepts, techniques, and facts across all areas of the specification. They use technical vocabulary and techniques with fluency, clearly demonstrating communication and numerical skills appropriate to a range of situations.
They demonstrate a good understanding of the relationships between data, evidence and scientific explanations and theories. They are aware of areas of uncertainty in scientific knowledge and explain how scientific theories can be changed by new evidence.
Candidates use and apply their knowledge and understanding in a range of tasks and situations. They use this knowledge, together with information from other sources, effectively in planning a scientific task, such as a practical procedure, testing an idea, answering a question, or solving a problem.
Candidates describe how, and why, decisions about uses of science are made in contexts familiar to them, and apply this knowledge to unfamiliar situations. They demonstrate good understanding of the benefits and risks of scientific advances, and identify ethical issues related to these.

They choose appropriate methods for collecting first-hand and secondary data, interpret and question data skilfully, and evaluate the methods they use. They carry out a range of practical tasks safely and skilfully, selecting and using equipment appropriately to make relevant and precise observations.
Candidates select a method of presenting data appropriate to the task. They draw and justify conclusions consistent with the evidence they have collected and suggest improvements to the methods used that would enable them to collect more valid and reliable evidence.

The qualifications regulators wish to make their publications widely accessible. Please contact us if you have any specific accessibility requirements.

First published in 2010.

Office of Qualifications and Examinations Regulation
Spring Place
Coventry Business Park
Herald Avenue
Coventry CV5 6UB
Telephone 0300 303 3344
Textphone 0300 303 3345
Helpline 0300 303 3346

www.ofqual.gov.uk