

Post 16

2001

# Five year review of standards

A level chemistry



Guarding standards

# Introduction

Every summer, the publication of GCSE and A level examination results prompts public interest in the standards of those examinations.

In 1996, Lord Dearing in his *Review of Qualifications for 16–19 Year Olds* made several recommendations to ensure that ‘there is a basis and accepted procedure ... for monitoring and safeguarding standards over time’. In the same year, SCAA (one of QCA’s predecessors) and the Office for Standards in Education jointly investigated standards in English, mathematics and science (chemistry) in 16+ and 18+ public examinations over time.<sup>1</sup>

The outcomes of this work were published in *Standards in Public Examinations 1975 to 1995*. One of the recommendations was 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. Physics, history, French and German should be included in the programme at an early stage.’

The five-yearly review of standards programme is a response to these recommendations. It is run by QCA in collaboration with the regulatory authorities for Wales and Northern Ireland, ACCAC and CCEA, and is designed to investigate the standards in A level and GCSE examinations. It aims to find out if:

the demand of syllabuses and their assessment instruments has changed over the last 20 years (examination demand);

the level of performance required of candidates at grade boundaries has changed over the last 20 years (grade standard).

Organised to run in five-year cycles, the programme was structured to cover every major subject during its first cycle. Each year, up to 100 independent specialists review around 2,000 exam scripts, drawn from all the awarding bodies, together with their associated syllabuses, question papers and mark schemes.<sup>2</sup>

---

<sup>1</sup> 16+ examinations cover GCE O level and Certificate of Secondary Education (up to 1987), and GCSE (from 1988).

<sup>2</sup> For the purposes of this report, the general term *awarding bodies* is used to cover both the A level examination boards and the GCSE examining groups.

# Methodology

Each study was organised in two stages:

- stage one – investigating changes in examination demand;
- stage two – investigating changes in standards of performance.

Each covered two sample years: the year of the study and 1995, the year used for the SCAA/Ofsted study.

## **Stage one: examination demand**

### ***Aim***

The aim of this review was to establish whether the demand of syllabuses and their assessment instruments changed over the period of the review.

### ***Evidence base***

The awarding bodies were asked to supply, for each subject, copies of one major syllabus from the most recent year. They were also asked to provide the related question papers, mark schemes, examiners' reports, and details of the procedures in operation at the time of each examination. The materials used in the SCAA/Ofsted study were available for comparison.

### ***The process***

A coordinator and three reviewers – independent experts from a variety of backgrounds – were appointed for each subject. Each coordinator was given a framework and asked to use it to describe the main differences between the syllabuses from the different years. This description was given to the reviewers, who were asked to study the syllabuses, question papers and mark schemes and independently judge whether the differences between years affected the demand of the examination. After the material had been reviewed, the team for each subject area met and discussed any issues. The coordinator then reported on the findings and identified any conclusions.

## **Stage two: standards of performance**

### ***Aim***

The aim of the second stage was to find out if the level of performance required of candidates at grade boundaries has changed over the period of the study. The review focused on the performance of candidates at grades A and E at A level, and grades A, C and, sometimes, F for 16+ examinations.

### ***Evidence base***

The awarding bodies were asked to provide 15 examples of candidates' work at the defined boundaries from the most recent year of examination. They were asked to

submit the complete examination work of candidates, including all examination papers, coursework and any oral examinations. The materials used in the SCAA/Ofsted study were available for comparison.

### ***The process***

A team of up to 12 reviewers was recruited for each subject. The reviewers came from a variety of backgrounds, including universities, selective and non-selective schools, maintained and independent schools, and further education institutions (including sixth form colleges). Some of them had backgrounds working for the various awarding bodies.

The coordinator from stage one was used again in this stage and the syllabus reviewers normally participated.

The review took place over two days. Before the meeting, each coordinator produced a general description of the standards expected for the grade boundaries in the study. Where these were available, published grade descriptions normally formed the basis of the performance descriptors. The coordinators were asked to take into account the fact that they would be looking at borderline performance rather than that comfortably in grade which is the intention of grade descriptions. The performance descriptors were discussed and agreed by the team at the start of the meeting.

Reviewers were each given a batch of scripts for a particular year, grade and awarding body. Working independently, they were asked to judge if the scripts matched the agreed grade description. They could categorise the work as:

- above the expected standard;
- slightly above the expected standard;
- at the expected standard;
- slightly below the expected standard;
- below the expected standard.

They were then given another batch of scripts of the same grade, either from another awarding body or of a different year from the same awarding body. They categorised these scripts and compared them with the first batch to identify any significant differences between candidates' performance. A sampling framework ensured adequate coverage of the sample. A copy of part of one framework is provided on page 4.

At the end of the two days, a plenary session was held and the reviewers discussed their findings and any significant issues. As with stage one, the coordinator reported on the findings and conclusions.

## **Limitations of the study**

Comparing examination standards over time is a complex task, heavily dependent on the evidence available and the ability of reviewers to make valid judgements on it. When considering the findings and conclusions, several limitations need to be kept in mind.

### ***Changes in syllabus and examination content***

Syllabuses and examination papers changed significantly over the period of the review. For example, in assessing GCSE science examinations, the three tiers of entry of 1995 had been reduced to two. Fundamental changes make it difficult for reviewers to make valid judgements about relative standards because they are not comparing like with like.

### ***Individual opinion***

Each individual places different values on each part of a subject. Agreed definitions of standards and frameworks show reviewers the standards they should work to, but it is difficult for them to avoid applying their own values. This can lead to differences in opinion about the same syllabus or piece of candidate's work.

### ***Lack of evidence***

While reviewers had syllabuses and examination papers (although not always mark schemes) for all the years in the study, they did not have all the evidence they needed to analyse standards of performance. This applies particularly to examination scripts. What was used in the SCAA/Ofsted study was work for separate components of the examination rather than the whole work of candidates. Coursework and any oral examinations were usually missing.

**Table 1: Sampling framework for part of a typical A level study**

**DAY 1**

8:30	BOARD A, GRADE A	BOARD A, GRADE E	BOARD F, GRADE A	BOARD F, GRADE E	BOARD C, GRADE A	BOARD C, GRADE E
10:00	1996 1-7	1996 1-7	1996 1-7	1996 7-1	1996 1-7	1996 15-8
10:10	BOARD A, GRADE A	BOARD A, GRADE E	BOARD F, GRADE E	BOARD F, GRADE A	BOARD C, GRADE A	BOARD C, GRADE E
11:30	1991 1-3	1991 1-3	1996 8-15	1996 7-1	1991 1-7	1991 15-8
11:50	BOARD A, GRADE E	BOARD A, GRADE A	BOARD C, GRADE E	BOARD C, GRADE A	BOARD E, GRADE A	BOARD D, GRADE A
1:05	1996 1-7	1996 15-8	1996 1-7	1996 8-15	1996 1-7	1996 15-8
2:15	BOARD A, GRADE E	BOARD A, GRADE A	BOARD A, GRADE E	BOARD B, GRADE E	BOARD E, GRADE E	BOARD D, GRADE E
3:30	1991 1-3	1991 3-1	1996 15-8	1996 15-8	1996 1-7	1996 15-8
3:30	BOARD B, GRADE A	BOARD D, GRADE E	BOARD B, GRADE A	BOARD D, GRADE E	BOARD D, GRADE A	BOARD E, GRADE A
4:45	1996 1-7	1996 1-7	1996 15-8	1991 4-1	1996 7-1	1996 8-15
5:05	BOARD B, GRADE E	BOARD D, GRADE E	BOARD B, GRADE E	BOARD D, GRADE E	BOARD D, GRADE E	BOARD E, GRADE A
6:20	1996 1-7	1991 1-4	1996 8-15	1986 4-1	1996 8-15	1991 1-3

**DAY 2**

8:30	BOARD C, GRADE E	BOARD E, GRADE E	BOARD E, GRADE A	EDEC , GRADE A 1996	BOARD F, GRADE A	BOARD A, GRADE E
9:45	1996 7-1	1996 15-8	1996 1-7	7-1	1996 8-15	1996 15-8
9:45	BOARD C, GRADE E	BOARD E, GRADE E	BOARD E, GRADE A	BOARD B, GRADE E	BOARD F, GRADE E	BOARD A, GRADE E
11:00	1991 1-7	1991 3-1	1991 3-1	1996 8-15	1996 8-15	1986 7-1
11:20	BOARD C, GRADE A	BOARD E, GRADE A	BOARD E, GRADE E	BOARD E, GRADE A	BOARD C, GRADE A	BOARD A, GRADE A
12:35	1996 7-1	1996 7-1	1996 8-15	1996 8-15	1996 15-8	1996 1-7
1:45	BOARD C, GRADE A	BOARD E, GRADE A	BOARD E, GRADE E	BOARD E, GRADE A	BOARD C, GRADE A	BOARD A, GRADE A
3:00	1991 7-1	1991 1-3	1991 1-3	1991 3-1	1991 15-8	1991 3-1

# A level chemistry: review of standards, 1995–99

## Introduction

SCAA, QCA's predecessor body, together with Ofsted, conducted an enquiry into examination standards. The results of the work, published in 1996 as *Standards in Public Examinations, 1975–1995* (SCAA, 1996), included a series of recommendations concerning future examinations in each subject reviewed. The subjects included A level chemistry, where changes were already in train at the time of the report.

The most significant of these changes were:

- a revision to the common subject core for first examination in 1996;
- a continuing move from linear to modular examinations;<sup>2</sup>
- the requirement for modular examinations to include a form of synoptic assessment to test candidates' understanding of the connections between different elements of the subject;
- a continuing trend away from practical examinations to internal assessment of practical work.

Syllabuses examined in 1999 were therefore those approved under the subject core revised prior to the 1996 work. Implementation of the 1996 report's recommendations has been effected for the specifications being taught from September 2000. This review offers, however, a chance to evaluate whether the last set of changes anticipated some of the recommendations or made them more urgent.

## Examination demand

### *Materials available*

Reviewers considered the syllabus documents, the question papers and associated mark schemes, and the examiners' reports for syllabuses from each of the awarding bodies in 1995 and 1999.

About 41,000 candidates took A level chemistry in 1999. About 54 per cent of those entered for the syllabuses used for that year in this study.

---

<sup>2</sup> The most popular A level chemistry assessment with all the awarding bodies involved in 1995 was a linear assessment whereas in 1999 in all cases it was a modular assessment. The review has therefore had to make comparisons between two very different forms of assessment.

### **Assessment objectives**

Between 1995 and 1999, changes to the assessment objectives (and aims) stated in the specifications were minor, generally amounting to redrafting with the intention of increasing clarity. However, AQA/N, Edexcel and OCR made changes to the way they grouped the assessment objectives when showing the weightings of these in the overall assessment. Meaningful comparisons of the weightings for these awarding bodies between 1995 and 1999, and between different awarding bodies in 1999, were therefore impossible.

### **Examining time**

It is noteworthy that, between 1995 and 1999, the time candidates had to answer written components was increased by all awarding bodies, significantly by Edexcel, OCR and particularly AQA/N. The number of written papers taken by a candidate to gain the A level qualification has also increased in all cases.

	Year	AQA/N	CCEA	EDEXCEL	OCR	WJEC
<b>number of papers</b>	1995	2	3	3	4	2
	1999	6	4	5	5	5
<b>total exam time/hours</b>	1995	5.0	5.75	5.5	5.5	6.0
	1999	9.0	6.0	7.5*	7.5	6.67

\* includes 10 minutes' reading time before the start of the synoptic paper.

In all cases, time spent in the examination room in 1999 had increased. This alone could be seen as increased demand. However, the work was examined in smaller sections and some candidates would not take all of the modules in the final sitting. The syllabus reviewers were of the opinion that, overall, these changes did not affect the demand of the assessments.

### **Practical work**

The table shows that the weightings awarded to practical work have changed in a variety of ways but were closer to one another in 1999 than in 1995. All 1999 practical assessments placed a greater emphasis on planning skills and this was judged to cause an increase in demand. During the period under consideration there was also a significant increase in the number of candidates entered for internal assessment of practical work.

	Year	AQA/N	CCEA	EDEXCEL	OCR	WJEC
<b>Theory/ Practical</b>	1995	80/20	85/15	80/20	80/20	85/15
<b>Theory/ Practical</b>	1999	82/18	80/20	80/20	83/17	80/20

### **Options**

The use of optional topics by the awarding bodies is summarised in the table below.



	Year	AQA/N	CCEA	EDEXCEL	OCR	WJEC
<b>Choice of topics studied</b>	1995	no choice	2 x 1 from 3	no choice	2 from 8	no choice
	1999	3 from 6	3 from 5	no choice	2 from 5 or 1 from 5 + 1 from 3	no choice

The AQA/N options and the OCR chemistry options in 1999 were considered comparable. Reviewers felt that the demand of the CCEA options had increased between 1995 and 1999 with the introduction of analytical and electro-chemistry but that the options collectively retained a high proportion of recall/knowledge.

### ***Subject content***

All awarding bodies made changes to the specified content in 1996. Comparison between 1995 and 1999 specifications shows that all awarding bodies except WJEC removed some inorganic chemistry which largely involved recall. Reviewers felt that for AQA/N and Edexcel, this did not adversely affect demand and still left realistic coverage of this area of chemistry. In OCR in 1999, however, the only compulsory group chemistry was a limited study of group I with study of groups II, IV and VII optional, leaving significantly less compulsory inorganic chemistry than in 1995. In CCEA in 1999, the only group which candidates had to study was group VII, although groups II and V could be studied as an option. Reviewers felt that to study only one group in the periodic table did not give a real insight into inorganic chemistry. However, for each awarding body it was considered that, when all the changes to the specified subject content were taken together, they represented no significant change in demand by comparison with 1995.

### ***Question papers***

The introduction of modular schemes has had a marked effect on the nature of the written assessments. In the linear 1995 assessments, all the awarding bodies used an objective question paper. In 1999, only CCEA used these questions, as part of three module tests, although WJEC had a section of short answer/multiple choice questions on each of two module papers. Reviewers agreed that removal of this type of question potentially reduced the syllabus coverage. In 1995, essay parts of questions had already become more structured and this trend has continued, leaving very few individual parts of any assessment worth more than five marks. In addition, Edexcel and CCEA 1995 assessments each contained a compulsory comprehension question but in 1999 this type of question only appeared in the WJEC synoptic paper. Reviewers considered that the removal of this type of question resulted in an increase in demand as some of the questions on comprehension passages were trivial and did not require any significant A level chemistry knowledge to answer them. The 1999 papers all had a high proportion of structured questions, contained within which were a limited number of short answer questions and some opportunities to write free response answers.

There was also a marked reduction in choice of questions, improving the fairness of the assessments.

	Year	AQA/N	CCEA	EDEXCEL	OCR	WJEC
Choice of question	1995	2 / 6	2 x 1 / 6	2 x 1 / 2 + 1 / 4	2 / 3 + 3 / 4 + 2 x 2 / 3 <sup>1</sup>	5 / 9
	1999	none	none <sup>1</sup>	2 / 4 <sup>2</sup>	5 x 1 / 2 <sup>1</sup>	none

<sup>1</sup> candidates answer all questions on the paper on option(s) they have studied.

<sup>2</sup> choice on one section of synoptic paper only.

Reviewers considered that the increase in the use of structured questions was more than balanced by several factors: the lack of question choice, the variety of question types within the questions set, and the opportunity to set some very testing questions focused on specific points. Reviewers decided that, overall, this change in the nature of the assessment represented an increase in demand. However, the 1999 CCEA assessment was considered comparatively heavy on recall, offering limited opportunities for candidates to demonstrate the higher skills of application and evaluation. Overall, calculations in the 1999 OCR assessment were judged too heavily structured.

### ***Synoptic assessment***

Reviewers looked at the synoptic element of the 1999 assessments because awarding bodies had been required to address this issue. The Edexcel 1999 synoptic paper was thought to be an excellent example of a meaningful way to assess synopticity. It was considered to be very well thought out, requiring candidates to draw knowledge from all modules and making use of the opportunity to set questions which could not be set within individual modules. WJEC also set a separate synoptic paper, but this was 40 minutes shorter and judged not so effective as a test of synopticity. Other awarding bodies included synoptic material from earlier modules in later modules but reviewers felt that this material was not as extensive nor as testing as that which appeared on the separate synoptic papers and that, by comparison with 1995 linear assessments, in which it was easy to set synoptic questions without the restriction of modules, demand in this area had fallen with the other three awarding bodies.

Questions involving social, economic and technological aspects did not feature prominently on the papers reviewed. It was felt that those questions which were set on this material were of a much higher demand in 1999 than in 1995, involving genuine A level chemistry.

In general, it was felt that while demand sometimes varied from module to module within a given assessment as a result of the different topics involved, each assessment package as a whole was entirely appropriate for advanced level.

### ***Mark schemes***

The clarity of mark schemes had generally improved from 1995 to 1999 making them more useful to teachers. However, reviewers were concerned that some aspects of some of the mark schemes might have affected demand. In 1999 CCEA and OCR mark schemes, some parts of questions had marking points in excess of the maximum marks for that part of the question. In some of these cases it was felt that this practice made it easier for candidates to score full marks in such parts than in

other parts of the paper. This was of particular concern when the question concerned was an optional one. All AQA/N and Edexcel papers reminded candidates of the necessity for good English and clear presentation. CCEA, OCR and WJEC papers went further in telling candidates that quality of language would be taken into account/assessed but, without detailed guidance on how these marks were awarded, it was not possible to assess how this might have affected demand. WJEC was the only awarding body to use half marks. In 1995 WJEC papers these were widely used and it was felt that demand on candidates was increased because most of these half marks would have been equivalent to a full mark on other awarding bodies' mark schemes. In the 1999 papers half marks were used less frequently but reviewers were still concerned that this practice increased demand.

## **Summary**

There had been a number of changes between 1995 and 1999, especially to the predominant approaches to assessment. These included a move to internal assessment of practical work, allowing a wider range of skills to be assessed. Most of the changes were considered neutral in effect on demand. In particular, the move from linear to modular examinations had not had an effect on demand. However, the question papers, although more structured in 1999 than in 1995, were considered slightly more demanding because of the reduction in levels of choice.

There were differences between awarding bodies in 1999. These were in the balance of skills assessed – with one, in particular, over-emphasising recall – and in the approach to synoptic assessment.<sup>3</sup> These were judged to lead to real differences in demand. Approaches to marking, especially the use of half marks, were also considered to lead to differences.

## **Performance at grade A and grade E**

### ***Materials available***

Reviewers considered candidates' work from all the awarding bodies in 1999 and materials from the 1996 enquiry for comparison. However, the data booklets used by OCR in 1995 and 1999 were missing and provision of evidence of practical skills was inconsistent: this made some of the judgements difficult or impossible.

The performance descriptions were developed from those in the new subject criteria, adjusted to reflect both changes in structure and content and the borderline nature of the candidates concerned. They were structured in four sections:

- skills shown;
- quantitative ability;
- synoptic ability;
- practical ability.

---

<sup>3</sup>The requirements for synoptic assessment have been greatly clarified and increased for the specifications for which teaching began in September 2000.

It was stressed that reviewers should not treat the descriptions as tick lists but should regard them as holistic descriptions of the qualities likely to be found in a borderline grade A or grade E.

### **Standards expected at grade A**

Candidates should be able to:

- recall and use knowledge from most parts of the specification;
- show a good understanding of the fundamental principles and concepts;
- show an ability to apply fundamental principles in new contexts;
- construct chemical equations (including redox half equations);
- use chemical equations quantitatively in a range of contexts;
- attempt most standard calculations in a logical manner, even when little guidance given;
- select relevant chemical knowledge in most situations;
- present their ideas clearly and logically, making correct use of chemical terminology;
- bring together and use knowledge and understanding from more than one area of the specification;
- formulate a plan for an experiment;
- show an understanding of techniques and safety required in experiments;
- recall experimental observations;
- interpret and explain experimental results.

### ***Performance at grade A***

In general, script reviewers considered that the work of borderline grade A candidates in 1999 matched the above description very well. Candidates showed the required level of knowledge, understanding and ability to apply their knowledge and a high level of ability in writing equations, carrying out calculations (even with little guidance) and drawing together material to present a logical, reasoned argument. These latter skills were considered to be tested very thoroughly in the Edexcel synoptic paper, which must have stretched even the most able grade A candidates. Undoubtedly candidates found this paper much more demanding than the other module tests and the paper was highly effective in providing discrimination. Reviewers had reservations about the ability of the more straightforward CCEA papers to discriminate adequately at this level, especially because the coursework produced consistently high marks for candidates, effectively adding little to the overall assessment. Responses suggested that the format of questions used by all awarding bodies was very effective in enabling candidates to show their skills and in both years candidates showed considerable communication skills.

Candidates' responses suggested that they found the AQA/N, Edexcel and OCR assessments slightly more demanding.

There was minimal testing of experimental work on theory papers and lack of evidence made generalisations on assessment of experimental work impossible. However, reviewers were impressed by the quality and variety of the 1999 coursework submitted by Edexcel and WJEC.

### ***Standards expected at grade E***

Candidates show an ability to:

- recall chemical knowledge from some parts of the specification;
- show some understanding of the fundamental principles and concepts;
- construct some simple chemical equations for straightforward, frequently-encountered reactions;
- make some progress in using simple equations quantitatively;
- carry out some straightforward calculations when guidance given;
- use some items of knowledge to answer structured questions;
- use some basic chemical terminology;
- apply knowledge and chemical principles contained within the specification to material presented in a familiar context;
- show some idea of how to set up an experiment;
- recognise risks in familiar experimental procedures;
- recall simple practical observations.

### ***Performance at grade E***

There were difficulties in using these (or any other) descriptions of grade E performance: many grade E candidates achieve the required total mark by achieving relatively high marks in some areas and low marks in others. However, reviewers judged that candidates had shown the appropriate standard to be awarded the grade. Candidates' knowledge tended to be very patchy; few showed real understanding of the principles and concepts; their equation writing was poor and they made little progress with questions testing synopticity. Generally, the standard of practical work seen from these candidates was comparatively high and many even showed ability to deal with planning aspects of this part of the assessment. Candidates at this level found it easier to deal with the more straightforward requirements of the CCEA assessment.

### ***Comparison of performance in 1995 and 1999***

Even though the linear assessments of 1995 were different from the modular assessments of 1999, reviewers felt that candidates had opportunities to show the same skills and did so. In 1995 assessments, candidates were able to show their knowledge through a paper of free response or structured essay questions but this same paper provided, in many cases, a wide choice of questions which enabled

candidates to avoid answering questions on areas of weakness. The 1999 structured questions provided better coverage and the design of these questions still provided candidates with demanding free response sections. Reviewers found the 1999 papers more user-friendly and accessible, providing candidates (particularly grade E candidates) with more opportunity to show what they knew and could do. In general, the 1999 papers were more searching: only in the case of CCEA at grade A was it felt that candidates showed a better performance standard in 1995 than in 1999.

## **Summary**

The standard of work at grade A and grade E in 1995 and 1999 was in line with expectations, and comparable both over time and, broadly, across awarding bodies.

# Key to the awarding bodies

During the period of the reviews, the number of awarding bodies operating fell. There are currently five: AQA/N, CCEA, Edexcel, OCR and WJEC. However, the three English awarding bodies came together through a number of mergers and a government requirement for unitary awarding bodies which could offer the range of GCSE, A level and GNVQ/VCE qualifications. This means that the qualifications used in the reviews came from a number of earlier examination boards and examining groups.

For the purposes of the reports the following abbreviations will be used:

AQA/N/A, AQA/N/N, CCEA, Edexcel, OCR and WJEC.

AQA/N/A covers AQA/N legacy A level syllabuses offered by AEB; legacy GCSE syllabuses offered by SEG; and O level syllabuses offered by AEB.

AQA/N/N covers AQA/N legacy A level syllabuses offered by NEAB, NEA and JMB; legacy GCSE syllabuses offered by NEAB and NEA; and O level syllabuses offered by JMB.

CCEA covers A level and GCSE syllabuses offered by CCEA, NISEAC and NISEC; and O level syllabuses offered by NISEC and NIGCEEB.

Edexcel covers A level and GCSE syllabuses offered by Edexcel, ULEAC and ULSEB; GCSE syllabuses offered by Edexcel, ULEAC and LEAG; and O level syllabuses offered by ULSEB.

OCR covers A level syllabuses offered by OCEAC, OCSEB, UCLES and UODLE; GCSE syllabuses offered by MEG; and O level syllabuses offered by OCSEB, UCLES and UODLE.

WJEC has retained the same name throughout the period.

© **Qualifications and Curriculum Authority (QCA) 2001**  
83 Piccadilly  
London W1J 8QA

**[www.qca.org.uk/](http://www.qca.org.uk/)**

***Order ref:*** QCA/01/765