Issue 16 June 2013

Research Matters

CAMBRIDGE ASSESSMENT

University of Cambridge
Local Examinations Syndicate

CAMBRIDGE ASSESSMENT

Citation

Articles in this publication should be cited as: Crisp, V. and Johnson, M. (2006). Examiners' annotations: Practice and purpose. *Research Matters: A Cambridge Assessment Publication*, 2, 10-12.

If you would like to comment on any of the articles in this issue, please contact Sylvia Green.
Email:
researchprogrammes@cambridgeassessment.org.uk

The full issue and previous issues are available on our website:
www.cambridgeassessment.org.uk/ca/Our_Services/Research

Research Matters : 16

A CAMBRIDGE ASSESSMENT PUBLICATION

Foreword

By virtue of having developed and managed qualifications for well over a century, Cambridge Assessment has accumulated considerable 'policy memory' regarding change in assessment and education. History tells us that change is neither a universal good nor a universal evil. Well-managed innovation has brought us assessments which measure with greater precision, have brought the benefits of education to a wider and more diverse range of candidates, and offer a better balance of cost and benefit. Whilst each innovation needs to be considered for the particular benefits it offers, 'change' needs to be considered in its own right, due to the challenges it presents and the pressures it places on individuals and systems. A period of change in qualifications and assessment inevitably reduces the capacity of education and training systems, since it directs effort away from 'normal' operation of the system and 'fine tuning' of existing qualifications. Continuous, ill-considered change in qualifications policy and regulation can lead to significant disruption in arrangements, reduction in confidence in qualifications, and unnecessary cost.

The UK has seen an extraordinary level of changes in the form, content and regulation of assessment – Nicky Rushton's article in this edition provides a fascinating overview of these revisions and 'system nudges'. In the face of this fast pace in change, calls for 'taking education out of politics' are increasingly heard. But both theory and practice tell us that it is impossible to 'take politics out of education', since education is intimately tied to social systems and political arrangements. It is, however, possible to take educational change out of 'the political cycle' – that is, a domestic electoral cycle which can, on occasion, encourage short-term thinking and ill-framed reform.

Without a flow of high quality evidence, 'evidence-based policy' is impossible. The studies included in this issue aim to ensure that this flow is sustained.

Tim Oates *Group Director, Assessment Research and Development*

Editorial

The articles in this issue address a range of investigations and perspectives on the theme of 'change'. Not only the details and descriptions of what has changed are addressed, but also how change can be best informed and what the intended and unintended consequences can be. In her article Rushton documents some of the changes that have occurred in secondary school education and general qualifications since 2000 and highlights the scale and scope of change over a relatively short period. Suto *et al.* consider change in the context of government policy and reform of GCE A levels in England, giving a chronological account of recent developments and reports on studies undertaken to research the views and experiences of stakeholders in schools, colleges and universities. The studies outlined exemplify ways in which a strong evidence base can be built to strengthen the development of qualifications.

Vidal Rodeiro discusses how attempts to widen the choice of qualifications in a subject can impact on the progression routes available to students. The analyses of national data reported here suggest that there is cause for concern in terms of opportunities afforded to certain groups of students according to the qualifications that they chose to take and the extent to which choices were made for them.

The next three articles illustrate how changes to assessment systems and structures can have both intended and unintended consequences. In the first of these Gill investigates patterns and effects of early certification in GCSEs in recent years. There has been concern that since the scrapping of Key Stage 3 tests candidates have been entered early and that this has meant that some have not achieved their potential. Added to the system change at Key Stage 3, it has been suggested that the pressure of the league tables and the accountability system has encouraged early entry to ensure achievement of Grade C which is a key performance measure. Gill also explores the impact of early entry at GCSE on uptake and performance at A level. The aspects of change highlighted in the work of Sutch and Wilson focus on the structure of A levels, the introduction of the new A* grade and the impact of both of these structural changes on the extent and outcomes of resitting by high performing A level candidates. This work provides interesting findings in an area that has been under researched while also correcting some widely held misconceptions about the extent of resitting at A level. The final article also explores the difficulty of common questions between tiered components to determine whether the tiered papers were functioning as intended. Dhawan and Wilson discuss possible reasons why some of the items may not have behaved as intended and they recommend ways of addressing some of the difficulties of tiering.

Sylvia Green *Director of Research*

Changing times, changing qualifications

Nicky Rushton Research Division

Introduction

During recent years there have been many changes in education and assessment in England. Since 2000, curricula have been updated, particular skills have been included then removed from assessment, several new qualifications for students in English secondary schools have been added to the Register of Regulated Qualifications and other qualifications have been withdrawn. When so many changes occur in a short space of time it is difficult to keep track of them, and the time at which they happened. This is particularly problematic when identifying the dates that qualifications were available.

This article tracks some of the changes that have occurred in England since 2000 in secondary school education and general qualifications. The year 2000 was chosen as the starting year for the analysis because it coincided with the first teaching of a new version of the National Curriculum in England and a major change to A levels. The article is divided into three sections: qualifications being added and withdrawn from the Register of Regulated Qualifications; changes to GCSEs (including the proposed English Baccalaureate Certificates); and changes to A levels.

For each section, a time line is included to provide an overview of the most important dates. Following on from this are summaries of the major events associated with each qualification/reform and references that have been identified for this information. These summaries do not give a comprehensive account of any particular event or document. Instead, relevant references are provided as a useful starting point to enable the reader to investigate further. Where possible, the references given are for published documents and publications, so that the reader can use them as a reference for the dates of particular events without needing to carry out further research. Occasionally it has been necessary to provide links to web pages instead. If this has been done, every effort has been made to ensure that the links are likely to prove permanent.

Introduction and withdrawal of qualifications

Since 2000, several new qualifications have been introduced and further qualifications have been withdrawn from the list of accredited qualifications. This section lists key dates and information for some of the qualifications offered in schools and colleges represented in Table 1 and the timeline in Figure 1.

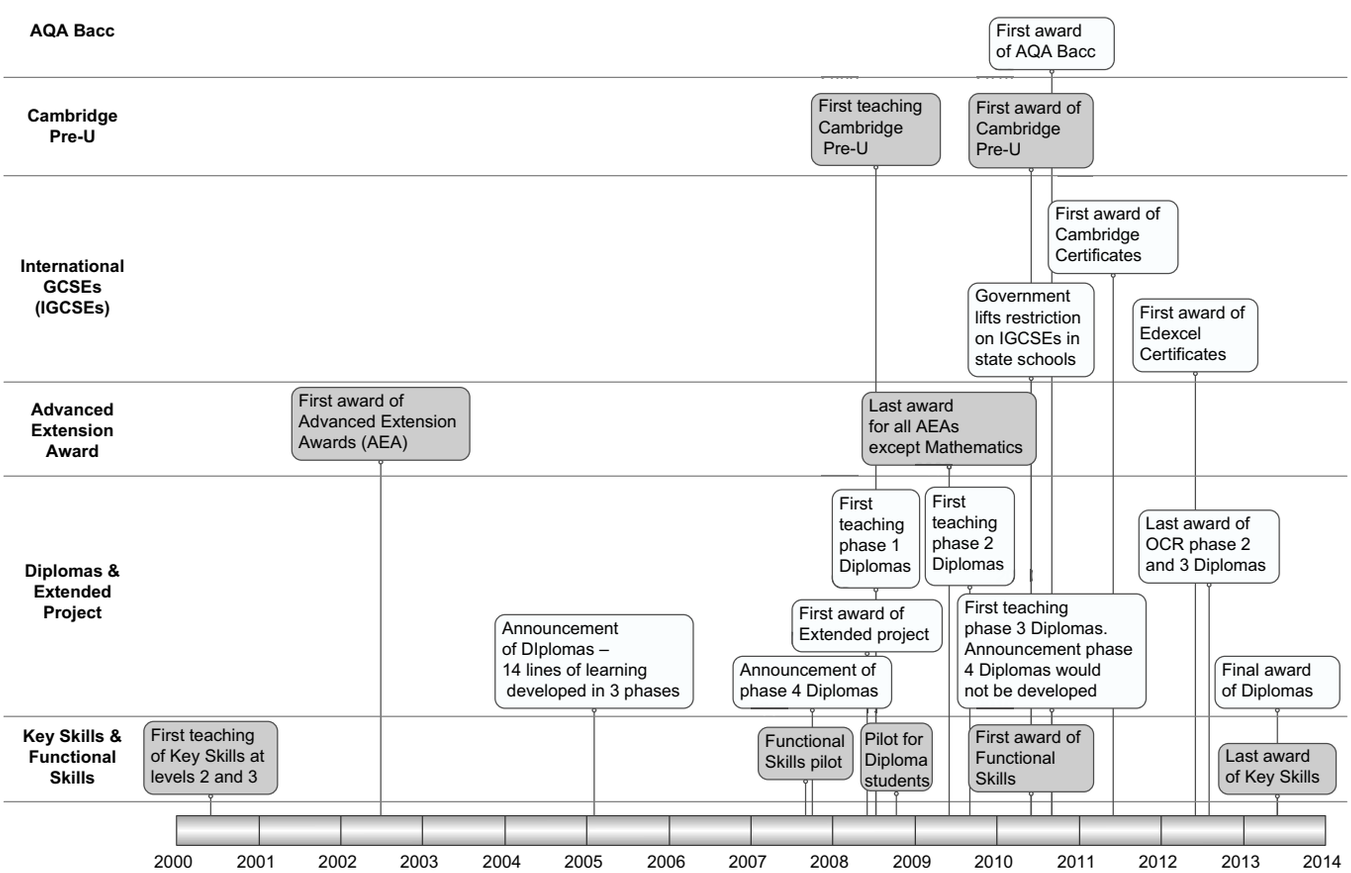


Figure 1: Key dates for offering qualifications

Table 1: Details of qualifications

Year	Details
Advanced Extension Award	
2002	The Advanced Extension Award (AEA) was introduced in 17 subjects, with the first examinations taking place in the summer of 2002 (UCAS, 2004). The specification[1] was based on the existing A level subject criteria and required no extra teaching. The examination was aimed at the top 10% of A level candidates and aimed to stretch them through more challenging questioning.
2009	All AEA examinations except Mathematics were withdrawn by the government, with the final examination taking place in the summer 2009 session. They were withdrawn as part of the changes to the A level specifications, which introduced stretch and challenge questions (QCDA, (2009).
On-going	Mathematics continues to be offered as an AEA award and has been accredited until 2015. No other AEAs are now available.
AQA Baccalaureate	
2008	The AQA Baccalaureate (AQA Bacc) was piloted. To qualify for the AQA Bacc, students must: gain an E or above in three A levels; take a further AS level in Citizenship, Critical Thinking, General Studies, Science in Society or World Development; complete an extended project qualification and undertake 100 or more hours of enrichment activities. (AQA, 2012).
2010	The AQA Bacc was awarded to any students meeting the criteria from 2010. (AQA, 2012).
Cambridge Pre-U	
2008	The Cambridge Pre-U was first available for schools to teach from September 2008. To qualify for the Cambridge Pre-U students must complete three Principal Subjects (A levels can be substituted for up to two subjects), undertake an Independent Research report and undertake a course in Global Perspectives. Each of these elements can also be undertaken as free-standing qualifications. (CIE, 2011a).
2009	The first students sat Cambridge Pre-U short course examinations in their Principal Learning subjects. Short courses were only available in Modern Foreign Languages for this session.
2010	The first (full course) students sat the Cambridge Pre-U examinations for their Principal Learning subjects.
Diplomas	
2005	A Diploma was first proposed in the Tomlinson report (DfES, 2004). In response the government announced the creation of Diplomas in 14 'lines of learning'[2] that would be available at levels 1, 2 and 3. The Diplomas were intended to prepare students for further education or employment by studying both specialised subject materials related to the 'lines of learning' (Principal Learning), and English, Mathematics and other subjects. The lines of learning were assigned to three phases for development, so that they would be introduced in three consecutive years. (DfES, 2005).
2007	The Government announced three new 'lines of learning' would be added to the Diploma. The new 'lines of learning' were: ● Science ● Humanities and Social Sciences ● Languages and International Communication. It was anticipated that the first teaching would take place in September 2011 (Ertl *et al*., 2009).
2008	The first teaching of the phase 1 Diplomas started in September. This included the following 'lines of learning' (and their related Principal Learning components): ● Construction and the Built Environment ● Creative and Media ● Engineering ● Information Technology ● Society, Health and Development. (Ertl *et al*., 2009).
2009	The first teaching of the phase 2 Diplomas started in September. This included the following 'lines of learning'(and their related Principal Learning components): ● Business, Administration and Finance ● Environmental and Land-Based Studies ● Hair and Beauty Studies ● Hospitality ● Manufacturing and Product Design. (Ertl *et al*., 2009).
2010	The first teaching of the phase 3 Diplomas started in September. This included the following 'lines of learning'(and their related Principal Learning components): ● Public Services ● Retail Business ● Sports and Active Leisure ● Travel and Tourism. (Ertl *et al*., 2009). In June 2010 the new Coalition Government announced that all development of the phase 4 diplomas would cease immediately. (DfE, 2010a).
2012	The last award of the phase 2 and 3 OCR Diplomas at levels 1 to 3 took place in summer 2012. This included the following 'lines of learning' (and their related Principal Learning components): ● Business, Administration and Finance ● Hospitality ● Manufacturing and Product Design ● Public Services ● Retail Business ● Sports and Active Leisure ● Travel and Tourism. (OCR, 2011a).
2013	The last award of Diplomas by any awarding organisation will take place in the summer of 2013 (Ofqual, 2010). This includes the 'lines of learning' for the following OCR phase 1 Diplomas and their related Principal Learning components (OCR, 2011a): ● Creative and Media ● Information Technology ● Society, Health and Development. Some aspects of Diplomas will continue, although the actual elements that will be continued and abandoned depend upon the individual awarding organisations. For example, OCR will still offer the Principal Learning in Engineering and the Foundation, Higher and Extended Project as stand alone qualifications (see below), and successful elements of other subjects which are also popular will be reflected in the development of future OCR qualifications (OCR, 2011b).
Extended Project	
2008	First teaching of the Extended Project. This requires students to project a report, dissertation, design, artefact or performance on a topic that they have chosen. It was introduced as a part of the Diploma, but was also available as a separate qualification in addition to A levels (OCR, 2008). The Extended Project will continue to be available after the final award of the Diploma (OCR, 2011b). OCR is also continuing Level 1 (Foundation) and Level 2 (higher) Project qualifications.
Functional Skills	
2007	The first pilot of Functional Skills started, with first teaching from September 2007. Functional Skills are practical skills of relevance to life, further education and employment in English, Mathematics and ICT. (Tribal, 2011).
2008	The Functional Skills pilot was extended in September to include all centres where students were studying for the Diploma. These students had to take functional skills as a compulsory part of their Diploma. (QCA, 2007a).
2010	First teaching of Functional Skills. This qualification is available to all students aged 14+ in English, Mathematics and ICT. (OCR 2010).

Table continues overleaf

1. A specification is "The document describing what will be assessed and how it will be assessed. Some awarding bodies use the more recent term 'specification' whilst others retain the traditional term 'syllabus'." (Elliott, 2011, p.11).
2. The lines of learning were: Construction and the Built Environment; Creative and Media; Engineering; Information Technology; Society, Health and Development; Business, Administration and Finance; Environmental and Land-Based Studies; Hair and Beauty Studies; Hospitality; Manufacturing and Product Design; Public Services; Retail Business; Sports and Active Leisure; Travel and Tourism.

Table 1: Details of qualifications *(continued)*

Year	*Details*
IGCSEs[3]	
2010	The Government lifted restrictions upon state schools offering International GCSEs. From September 2010 state schools were able to offer accredited International GCSE courses alongside GCSE courses. (DfE, 2010a).
2011	The first state school students sat CIE's Cambridge IGCSEs (accredited by Ofqual as 'Cambridge International Level 1/Level 2 Certificates') in the summer 2011 examination session (CIE, 2011b). (Unaccredited Cambridge IGCSEs, which keep their name, continue to be offered to overseas schools and non-state schools in the UK.)
2012	The first state school students sat Edexcel International GCSEs (accredited by Ofqual as 'Edexcel Level 1/Level 2 Certificates') in the summer 2012 examination session. (Edexcel, 2012).
Key Skills	
2000	Key Skills assessed students' achievement in three skills valued by employers and higher education: communication, application of number, and information technology. The first teaching of Key Skills at levels 2 and 3 took place in 2000. (QCA, 1999).
2013	The last certification of Key Skills will take place by the end of September 2013. (OCR, 2012a).

3. There may be some confusion about the IGCSE and the International Certificate qualifications. IGCSEs were developed by CIE and Edexcel as the International version of GCSEs. The IGCSEs offered by CIE are called Cambridge IGCSEs, whilst those offered by Edexcel are International GCSEs. International Certificates are the accredited versions of the International GCSEs. They appear on the Register of Regulated Qualifications as Cambridge International Certificates if they are offered by CIE and Edexcel Certificates if they are offered by Edexcel. Whilst they are often informally referred to as IGCSEs or International GCSEs, any documentation falling under the remit of Ofqual (e.g. results certificates) cannot use these titles.

Changes to GCSEs

There have been many changes to GCSEs since 2000, as detailed in Table 2 and the timeline in Figure 2. For the first few years these were limited to changes that were introduced with the introduction of new specifications. Since the beginning of 2009 the changes have been rather more frequent, and have been implemented for a number of reasons. Some changes, such as the revisions of the science specifications, have followed recommendations arising from Ofqual's qualification monitoring programme. Others, such as the change back to linear specifications, have been implemented as a result of the policy changes following the change of government. As the Coalition Government has made so many changes, they have been reported in their own timeline (Figure 3) to make it easier to follow when each change was made. In addition, as the changes related to the replacement of GCSEs with English Baccalaureate Certificates (EBCs) overlapped with several earlier changes, the details of these changes are reported in their own table (Table 3).

Table 2: Details of GCSE changes

Subject	*Details*
2001	
Most subjects	First teaching of new specifications for most GCSEs.[4]
2002	
Applied Art & Design, Applied Business, Applied ICT, Applied Science, Engineering, English, English Literature, Health & Social Care, Leisure & Tourism, Manufacturing, Psychology	First teaching of new GCSE specifications in remaining subjects (see left).[4]
2003	
Most subjects	First examination of new specifications for most GCSEs.[4]
2004	
See 2002 list	First examination of new GCSE specifications in remaining subjects (see 2002 subject list).[4]
2006	
Science	First teaching of new Science GCSEs. Includes Twenty First Century Science suite.[4]
2009	
Most subjects	First teaching of new specifications. Changes included: ● Controlled assessment replaced coursework ● Most courses became modular ● A requirement was added that 40% of assessment is terminal (taken at end of course) (QCA, 2009).
Science specifications	Ofqual (2009a) reported on monitoring new Science from the first full examination session. Variation between boards and concerns about reliability and validity lead to the recommendation that criteria were redeveloped and new specs drawn up in response to this.
2010	
English, Mathematics, ICT	First teaching of new specifications for English, Mathematics and ICT. (QCA, 2009).
N/A	In November, the White Paper 'The Importance of Teaching' (DfE, 2010b) was published. It contained the coalition policies for education. Of importance to GCSE were announcements that the re-sit rules would change; linear exams would be re-introduced; and spelling, punctuation and grammar (SPAG) would be given more importance.
2011	
Most subjects	The first certification of the new modular single and double award GCSEs with a 40% terminal requirement and controlled assessment instead of coursework. (QCA, 2009).

Table continues on page 6

4. Note: No direct reference is available for this. Dates were taken from the books of specifications from that examination year.

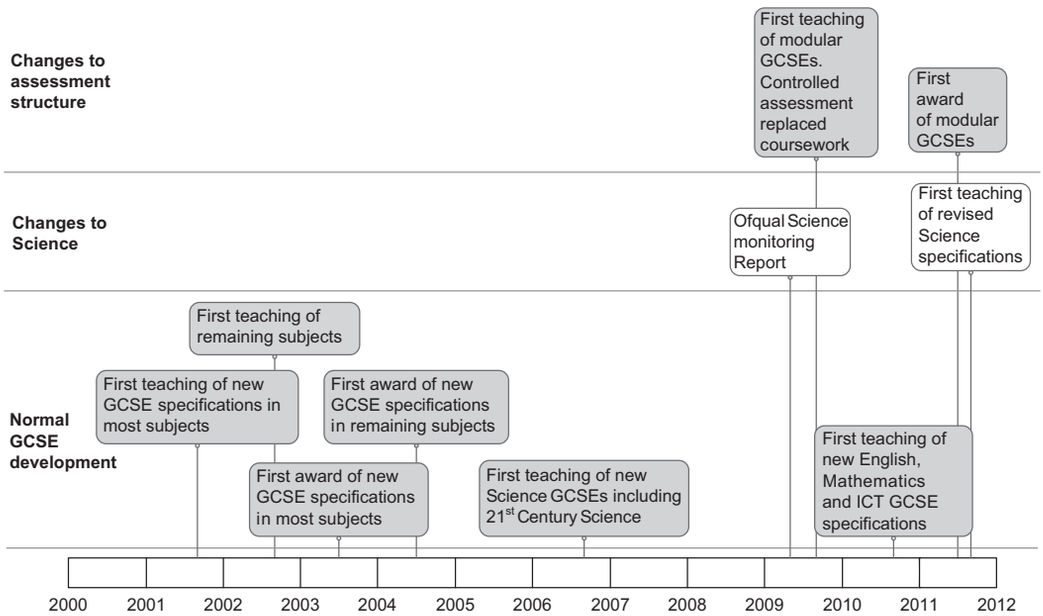


Figure 2: Changes to GCSEs prior to Coalition Government

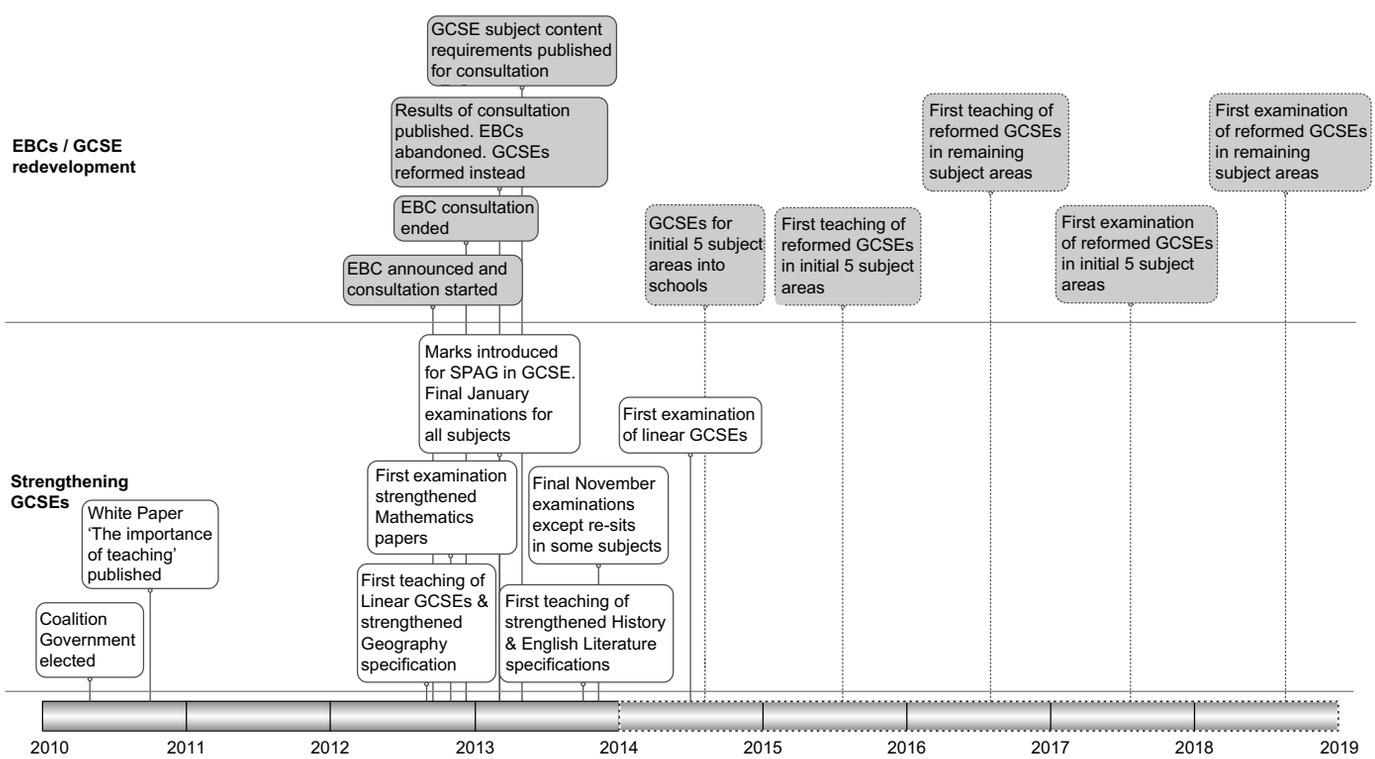


Figure 3: Changes to GCSEs since the election of the Coalition Government

Table 2: Details of GCSE changes *(continued)*

Subject	*Details*
2011 *(continued)*	
Science	The first teaching of the new science GCSEs, altered in response to the Ofqual comparability report. (Ofqual, 2011).
N/A	The *Daily Telegraph* published their report into the exam seminars run by the awarding organisations. (*The Telegraph*, 2011).
2012	
All subjects	First teaching of linear GCSE specifications in September 2012. (OCR, 2012b).
Geography, English Literature, History and Mathematics	After the 2011 *Daily Telegraph* investigation into examination seminars (*The Telegraph*, 2011), Ofqual announced that these four subjects would be strengthened. Changes were announced for the papers in Mathematics and the specifications of the other subjects. (Ofqual, 2012a).
Geography	First teaching of the strengthened Geography specification from September 2012. (Ofqual, 2012b).
Mathematics	First examination of the strengthened Mathematics papers in the November session (the specification remained unchanged from 2010). (Ofqual, 2012b).
2013	
English, English Literature, Geography, History, RE	Increased emphasis on spelling, punctuation and grammar with extra marks awarded for this from the January examination session onwards. (OCR, 2012b).
All subjects	Final January GCSE session. January examinations will only be available to candidates who will certificate in 2013. (OCR, 2012b).
English Literature, History	Projected first teaching date for strengthened specifications from September 2013. (Ofqual, 2012b).
Most subjects	Final November session for most subjects. November examinations will only be available to students who will certificate in 2013. Future November sessions will only be available for English, English Literature and Mathematics, and students will have to re-take the whole qualification rather than individual units/examinations. (OCR, 2012b).
2014	
All subjects	First certification of linear GCSEs in June examination session. (OCR, 2012b).

Table 3: Changes to GCSEs since the election of the Coalition Government

Details
2012
On the 17th September, the Government announced the development of new qualifications, EBCs, to replace GCSEs. Initially EBCs were intended to be available for six subjects: English, Mathematics, Sciences, History, Geography and Languages. It was proposed that each subject should be examined by one awarding organisation, (market reform) following a franchised model, and that all the assessment should be through examinations at the end of the course. The qualifications would be designed so that they were suitable for the full range of candidates taking GCSEs, but there would not be any tiering[5] used in the examination papers. It was also proposed that a new grading structure would be developed for the qualifications, but was not revealed in the initial announcement. (DfE, 2012a). A consultation on the proposed changes ran from the 18th September until the 10th December 2012. (DfE, 2012b).
2013
On the 7th February, the results of the consultation were announced. The proposed EBCs would not be developed and the market reform involving the franchised model would be abandoned. There would be major reforms to the existing GCSEs, retaining many of the changes initially proposed for EBCs. It was proposed that the reformed GCSEs would be linear, with more stretching assessments. Although tiered papers would still be abandoned, it was suggested that core and extension papers may be allowed if all students could be entered for them. The changes would be made to the first five subject areas[6] so that they were ready for first teaching in 2015. (House of Commons, 2013).

Details
2014
Projected date for specifications in initial five subject areas being available in schools. (Ofqual, 2013).
2015
Projected first teaching of new GCSEs in the initial five subject areas. (House of Commons, 2013).
2016
Projected first teaching of new GCSEs in the remaining subject areas. (DfE, 2013a).
2017
Projected first examination of new GCSEs in the initial subject areas.
2018
Projected first examination of new GCSEs in the remaining subject areas.

5. Currently, GCSE papers have two tiers: Foundation and Higher. Students can only be entered for one tier, and the range of grades that they can achieve depends upon the papers that they are entered for.

6. The initial five subject areas will be: English, Mathematics, the Sciences, History and Geography.

Changes to A levels

A levels have undergone fewer changes than GCSEs over the years since 2000. Like GCSEs, several of the changes have been as the result of government policy changes. As with GCSEs, the changes (represented in Table 4 and in the timeline in Figure 4) generally affect all subjects.

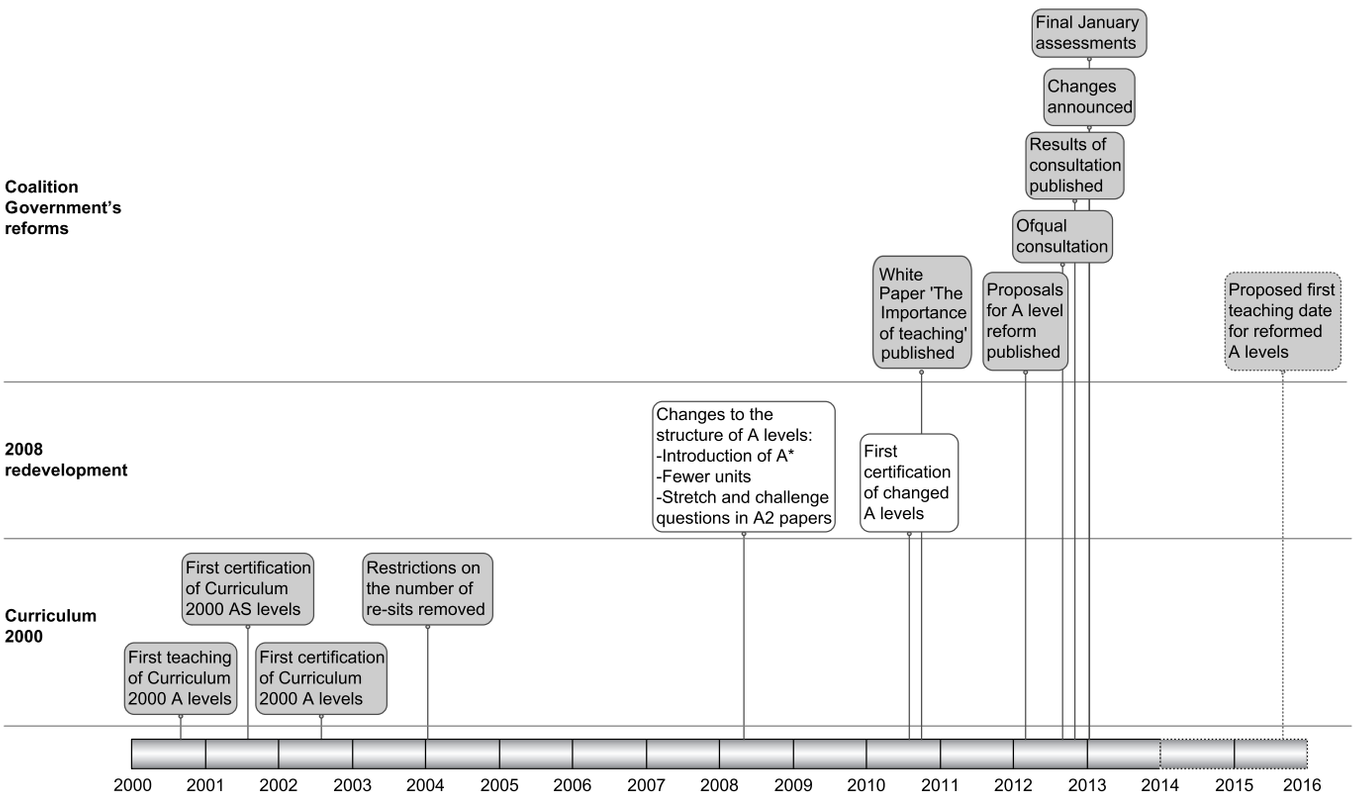


Figure 4: Changes to A levels since 2000

Table 4: Details of A level changes

Details
2000
First teaching of Curriculum 2000.
AS levels were introduced as a half-way stage for all[7] A levels and were worth 50% of the final A level mark.
The new specifications were divided into modules, with an examination at the end of each module. Most subjects consisted of six modules: three at AS level and three at A2 level. Students could either sit these examinations as they went along (a modular course) or sit them all at the end (a linear course).
Students were able to re-sit each module once to improve their marks.
The modular nature of the course and the inclusion of the AS level meant that students were expected to take more subjects in the first year of their courses, dropping one or two to specialise in the second year. This was done to bring England in line with the courses offered in European countries. (Ofsted, 2003).
2001
First certification of the new AS levels following the introduction of Curriculum 2000.
2002
First certification of the Curriculum 2000 A levels.

Details
2004
Following recommendations made by Tomlinson (2002) in the report into A level grading in 2002, restrictions on re-sits were removed from the January 2004 session onwards. This meant that students were able to re-sit each module an unlimited number of times. (Poon Scott, 2012).
2008
A levels were changed for first teaching in September 2008.
The units for most subjects were reduced from six to four, with two units for AS and two units for A2. Science, Mathematics and Music kept six units.
Stretch and challenge questions and synoptic assessment were introduced in the units for A2.
The A* grade was added to the A level grade (but not AS) for students who had achieved a grade A at A level and had achieved 90% or more of the UMS[8] marks on their A2 units.
(QCA, 2007b).

Table continues on page 8

7. Before Curriculum 2000, AS levels were usually stand-alone qualifications which were intended to have half the breadth and depth of an A level. However, in certain subjects which were assessed modularly (for example, Mathematics and Science) some of the modules could contibute to either an AS or an A level.
8. The uniform mark scale (UMS) is a scale where the range of marks for a particular A level grade is identical, regardless of the subject, paper or year. The marks awarded on individual A level assessments are converted onto the UMS scale to allow the marks on each assessment to be added up and a final grade to be calculated. (Ofqual, 2009b).

Table 4: Details of A level changes *(continued)*

Details
2010
The first certification of the new A levels. In November, the White Paper 'The Importance of Teaching' (DfE, 2010b) was published. It contained the coalition policies for education. Of importance to A levels were announcements that: ● Ways of involving universities in A level development would be explored ● Assessments would be modified to contain deep synoptic learning ● The re-sitting rules would be evaluated and possibly changed.
2012
In April, proposals for a further reform of A level were published. The proposals arose from the government's education White Paper (DfE, 2010b) and were a response to concerns that A levels were not good preparation for undergraduate study. The proposals were: ● For universities to be involved in the design and development of A levels ● To consider whether the division of A levels into AS and A2 should continue ● To consider whether January re-sits should be allowed. (DfE, 2012c). Ofqual held a consultation on the proposed reforms between June and September 2012 (Smith, Mitchell and Grant, 2012). The findings from the consultation were published in November 2012 and Ofqual announced that there would be no more January examination sessions after September 2013 (Ofqual, 2012c).
2013
Final January GCSE session for all candidates. (Ofqual, 2012c). In late January the following changes to A levels were announced by the Government (Gove, 2013): ● An advisory group, consisting of representatives from the Russell Group universities, would advise Ofqual on the content of A levels ● AS level would be retained, but as a standalone qualification which was at the same level as A levels, rather than as a part of A levels ● A levels would be made fully linear.
2015
September 2015 is the proposed first teaching date for reformed A levels. (DfE, 2013b).

References

Introduction and withdrawal of qualifications

AQA (2012). *AQA Level 3 Baccalaureate Specification: Guide to the AQA Bacc*. Available online at: http://filestore.aqa.org.uk/subjects/AQA-BACC-W-SP-10.PDF (Accessed 26 April 2013).

CIE (2011a). *Cambridge Pre-U: A guide for schools*. Available online at: http://www.cie.org.uk/docs/qualifications/preu/guide/Cambridge%20Pre-U%20Guide%20for%20Schools.pdf (Accessed 26 April 2013).

CIE (2011b). *Cambridge IGCSE students receive results today.* Press release. Available online at: http://www.cie.org.uk/news/pressreleases/detail?pressrelease_id=40129 (Accessed 26 April 2013).

Department for Education (DfE) (2010a). *Government announces changes to qualifications and the curriculum*. London: DfE. Available online at: https://www.gov.uk/government/news/government-announces-changes-to-qualifications-and-the-curriculum (Accessed 26 April 2013).

Department for Education and Skills (DfES) (2004). *14–19 Curriculum and Qualifications Reform: Final Report of the Working Group on 14–19 Reform*. Annesley: DfES Publications. Available online at: http://dera.ioe.ac.uk/id/eprint/11961 (Accessed 26 April 2013).

Department for Education and Skills (DfES) (2005). *14–19 Education and Skills White Paper*. London: DfES. Available online at: http://www.educationengland.org.uk/documents/pdfs/2005-white-paper-14-19-education-and-skills.pdf (Accessed 26 April 2013).

Edexcel (2012). *The Edexcel Certificate: Mathematics*. Available online at: http://www.edexcel.com/quals/igcse/edexcel-certificate/maths/Pages/default.aspx (Accessed 26 April 2013).

Elliott, G. (2011). A guide to comparability terminology and methods. *Research Matters: A Cambridge Assessment Publication*, Special Issue **2**, 9–19. Available online at: http://www.cambridgeassessment.org.uk/ca/digitalAssets/197994_RM_Special_Issue_2__web_.pdf (Accessed 26 April 2013).

Ertl, H., Stanley, S., Huddleston, P., Stasz, C., Laczik. A, and Hayward, G. (2009) *Reviewing Diploma Development: Evaluation of the Design of the Diploma Qualification*. Nottingham: Department for Children, Schools and Families. Available online at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/189666/DCSF-RW080.pdf.pdf (Accessed 26 April 2013).

OCR (2008). *Extended Project: Level 3 Extended Project.* OCR. Available online at: http://www.ocr.org.uk/images/80773-centre-handbook-specification.pdf (Accessed 26 April 2013).

OCR (2010). *Functional Skills: Deliver that spark.* OCR. Available online at: http://www.ocr.org.uk/images/75392-functional-skills-brochure.pdf (Accessed 26 April 2013).

OCR (2011a). *Notice to centres: Withdrawal of Diploma lines of learning (Principal Learning)*. Available online at: http://www.ocr.org.uk/images/105417-notice-to-centres-diploma-and-principal-learning-provision.pdf (Accessed 26 April 2013).

OCR (2011b). *Announcement on Diploma (Principal Learning) provision*. Available online at: http://www.ocr.org.uk/news/view/announcement-on-diploma-principal-learning-provision/ (Accessed 26 April 2013).

OCR (2012a). *OCR Key and Basic Skills – Guide to the end of entry arrangements*. Available online at: http://www.ocr.org.uk/images/75945-key-and-basic-skills-end-of-entry-arrangements.pdf (Accessed 26 April 2013).

Ofqual (2010). *The Diploma.* Available online at: http://www.ofqual.gov.uk/qualifications-assessments/89-articles/16-diploma (Accessed 26 April 2013).

QCA (1999). *Curriculum guidance for 2000: Implementing the changes to 16–19 qualifications.* Coventry: Qualifications and Curriculum Authority. Available online at: http://homepages.shu.ac.uk/~edsjlc/ict/qca/a-as_new_2000/framework/curric-implement.pdf (Accessed 26 April 2013).

QCA (2007a). *Functional skills: essential for life, learning and work*. Coventry: Qualifications and Curriculum Authority. Available online at: http://www2.ofqual.gov.uk/downloads/category/129-guidance?download=203%3Aa-guide-to-functional-skills (Accessed 26 April 2013).

QCDA (2009) *Advanced Extension Awards.* Coventry: Qualifications and Curriculum Development Agency. Available online at: http://webarchive.nationalarchives.gov.uk/20100210222040/qcda.gov.uk/4115.aspx (Accessed 26 April 2013).

Tribal (2011). *Evaluation of the Functional Skills Pilot: Summative Report*. Coventry: Qualifications and Curriculum Development Agency. Available online at: http://www.great-learning.co.uk/wp-content/uploads/2011/03/Evaluation_of_the_Functional_Skills_Pilot_Summative_Report-feb-11.pdf, (Accessed 26 April 2013).

UCAS (2004). *Expert group report for awards seeking admission to the UCAS Tariff: Advanced Extension Awards.* UCAS. Available online at: http://www.ucas.com/documents/tariff/tariff_reports/aeareport.pdf (Accessed 26 April 2013).

Changes to GCSEs

Department for Education (DfE) (2010b). *The Importance of Teaching.* Norwich: TSO. Available online at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/175429/CM-7980.pdf (Accessed 26 April 2013).

Department for Education (DfE) (2012a). *Reforming Key Stage 4 Qualifications*. Available online at: http://dera.ioe.ac.uk/id/eprint/15560 (Accessed 26 April 2013).

Department for Education (DfE) (2012b). *Reforming Key Stage 4 Qualifications: Consultation Response Form.* Available online at: http://dera.ioe.ac.uk/id/eprint/15560 (Accessed 26 April 2013).

Department for Education (DfE) (2013a). *Letter of 6th March 2013 from the Secretary of State for Education to the Chief Executive of Ofqual*. Available online at: http://www.ofqual.gov.uk/files/2013-02-07-letter-from-michael-gove-reform-of-ks4-qualifications.pdf (Accessed 26 April 2013).

House of Commons (2013). Hansard, 7 February 2013, Column 441–443 on *Curriculum and Exam Reform*. Available online at: http://www.publications.parliament.uk/pa/cm201213/cmhansrd/cm130207/debtext/130207-0001.htm (Accessed 26 April 2013).

OCR (2012b). *GCSE changes – Linear assessment and SPAG: Time line for changes – key dates*. OCR. Available online at: http://www.ocr.org.uk/images/16262-gcse-changes-linear-assessment-and-spag.pdf (Accessed 26 April 2013).

Ofqual (2009a). *The new GCSE science examinations: Findings from the monitoring of the new GCSE science specifications: 2007 to 2008.* Office of Qualifications and Examinations Regulation. Available online at: http://www.ofqual.gov.uk/files/ofqual-09-4148_GCSE_science_2007_2008_report.pdf (Accessed 26 April 2013).

Ofqual (2011). *GCSE science accreditation.* Office of Qualifications and Examinations Regulation. Available online at: http://www2.ofqual.gov.uk/news-and-announcements/83/549 (Accessed 26 April 2012).

Ofqual (2012a). *Ofqual confirms changes to GCSEs.* Office of Qualifications and Examinations Regulation. Available online at: http://www.ofqual.gov.uk/news-and-announcements/130-news-and-announcements-press-releases/820-ofqual-confirms-changes-to-gcses (Accessed 26 April 2013).

Ofqual (2012b). *Announcement on 'tightening' of GCSEs in key subjects*. Office of Qualifications and Examinations Regulation. Available online at: http://www.ofqual.gov.uk/news-and-announcements/83/854 (Accessed 26 April 2013).

Ofqual (2013). *Letter of 6th February 2013 from the Chief Executive of Ofqual to the Secretary of State for Education*. Available online at: http://www.ofqual.gov.uk/files/2013-02-07-letter-to-michael-gove-key-stage-4-qualifications-reform.pdf (Accessed 26 April 2013).

QCA (2009). *Information for teachers: changes to GCSEs: including controlled assessment.* Coventry: Qualifications and Curriculum Authority. Available online at: http://dera.ioe.ac.uk/id/eprint/9280 (Accessed 26 April 2013).

The Telegraph (2011). Exams that fail the test. *The Daily Telegraph*, 7th December 2011. Available online at: http://www.telegraph.co.uk/comment/telegraph-view/8940739/Exams-that-fail-the-test.html (Accessed 26 April 2013).

Changes to A levels

Department for Education (DfE) (2010b). *The Importance of Teaching: The Schools White Paper 2010*. Norwich, TSO. Available online at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/175429/CM-7980.pdf (Accessed 26 April 2013).

Department for Education (DfE) (2012c). *Letter of 30th March 2013 from the Secretary of State for Education to the Chief Executive of Ofqual on the reform of GCE A levels*. Available online at: http://www2.ofqual.gov.uk/files/2012-03-31-michael-gove-letter-to-glenys-stacey.pdf (Accessed 26 April 2013).

Department for Education (DfE) (2013b). *Letter of 22nd January 2013 from the Secretary of State for Education to the Chief Executive of Ofqual on the reform of GCE A levels*. Available online at: http://www.ofqual.gov.uk/files/24-01-2013-ofqual-letter-reform-of-gcse-a-levels.pdf (Accessed 26 April 2013).

Ofsted (2003). *Curriculum 2000: implementation*. Ofsted. Available online at: http://www.ofsted.gov.uk/sites/default/files/documents/surveys-and-good-practice/c/Curriculum%202000%20implementation%20%28PDF%20format%29.pdf (Accessed 26 April 2013).

Ofqual (2009b). *A levels: The official student guide to the system*. Office of Qualifications and Examinations Regulation. Available online at: http://www2.ofqual.gov.uk/files/A-levelGuide.pdf (Accessed 26 April 2013).

Ofqual (2012c). *Ofqual announces changes to A levels*. Press release. Office of Qualifications and Examinations Regulation. Available online at: http://www.ofqual.gov.uk/news/ofqual-announces-changes-to-a-levels/ (Accessed 26 April 2013).

Poon Scott, E. (2012). Short-term gain at long-term cost? How resit policy can affect student learning. *Assessment in Education: Principles, Policy and Practice*, **19**, 4. Available online at: http://dx.doi.org/10.1080/0969594X.2012.714741 (Accessed 26 April 2013).

QCA (2007b). *Changes to A levels: Improving the assessment system.* Coventry: Qualifications and Curriculum Authority. Available online at: http://dera.ioe.ac.uk/id/eprint/7171 (Accessed 26 April 2013).

Smith, J., Mitchell, T. & Grant, A. (2012). *Analysis of the consultation carried out into higher education involvement in GCE A levels and amended GCE A level criteria (design rules)*. Coventry: Qualifications and Curriculum Authority. Available online at: http://www.ofqual.gov.uk/files/2012-11-07-Analysis-of-the-consultation-into-he-involvement-in-a-levels-and-amended-criteria.pdf (Accessed 26 April 2013).

Tomlinson, M. (2002). *Inquiry into A Level Standards: Final Report.* London: Department for Education and Skills. Available online at: https://www.education.gov.uk/publications/eOrderingDownload/Tomlinson%20Final%20%20Report.pdf (Accessed 26 April 2013).

A level reform: Is the Government in tune with its stakeholders?

Irenka Suto, Sanjana Mehta, Simon Child, Frances Wilson Research Division and **Elizabeth Jeffery** Independent Research Psychologist

Introduction

Reformed GCE A levels are on the educational horizon for many students and their teachers. Awarding bodies are in the process of redeveloping their courses and from September 2015, the new syllabuses will be taught in sixth forms across England. In this article, we give a chronological account of the recent developments in Government policy which have fed into these plans. Alongside this account, we describe five studies that we have undertaken within our Higher Education (HE) Engagement research programme. An overarching aim of our research has been to ascertain the views and experiences of stakeholders in schools, colleges and universities on multiple aspects of A level reform.

Early Coalition Government policy

In November 2010, the Department for Education (DfE) published a White Paper – *The importance of teaching* – in which it signalled its intentions for whole-system reform in England. The paper set out the (then) recently elected Coalition Government's early views and concerns relating to a range of educational issues, including teaching and leadership, behaviour, accountability, school improvement, and funding.

Within a chapter of the paper on curriculum, assessment and qualifications, the Government indicated its desire to reform A levels. Moreover, it made clear its commitment to engage Higher Education Institutions (HEIs) in the reform process:

> *A levels are a crucial way that universities select candidates for their courses, so it is important that these qualifications meet the needs of higher education institutions. To ensure that they support progression to further education, higher education or employment, we are working with Ofqual, the awarding organisations and higher education institutions to ensure universities and learned bodies can be fully involved in their development.* (Department for Education, 2010, Section 4.47)

The idea that universities should play a role in developing A levels was not a new one. In the 1950s, 60s and 70s, lecturers and other university staff worked extensively with examination boards to ensure that A levels prepared students for higher level study (Kingdon, 1991; Raban, 2008). Arguably, the grounds for this collaboration have since strengthened, as the number of students taking A levels in preparation for university has risen considerably; 52% of all A level students now progress directly to HE (DfE, 2012b). On the contrary, however, the influence of HE on the design and content of A levels waned significantly in the last two decades. Successive governments and government-funded bodies became more actively involved instead.

Stakeholder re-engagement within Cambridge Assessment

Prior to the publication of the Government's White Paper (DfE, 2010), Cambridge Assessment was already attempting to reassert a more balanced ecosystem in which it worked with universities as well as schools and colleges. Five years ago, for example, Cambridge International Examinations (our international awarding body) ensured that HE representatives were successfully involved in the development of the Cambridge Pre-U, an alternative academic qualification for 16 to 19 year olds. In early 2010, Cambridge Assessment launched a consultation paper on the future of A levels and staff engaged in extensive discussions on the topic with personnel at the Department for Education. Alongside this work, OCR (our UK awarding body) set up a series of consultative forums as part of a wide-ranging stakeholder re-engagement programme. The forums' dual aims were: (i) to provide regular opportunities to update stakeholders on relevant developments in qualifications policy and practice; and (ii) to obtain their views on such developments, in order to inform OCR's plans.

The consultative forums are on-going and their memberships continue to grow. There are currently 11 subject-specific forums, each of which meets twice a year. Membership comprises a range of people from across the educational community, including HE lecturers and admissions tutors, teachers, employers, and representatives from subject and professional associations, learned societies, and charitable organisations. Over 70 different institutions are involved. Additionally, an HE strategic forum meets three times a year and its members are drawn solely from HEIs.

Throughout 2011, the merits and weaknesses of A levels were discussed by over 240 interested forum members, around a hundred of whom were HE representatives. Many lecturers and tutors made suggestions as to how curricula and associated assessments for 16 to 19 year olds could be improved in the future. This engendered a need to gauge the wider representativeness of the views expressed, to prioritise concerns, and to conduct some more detailed investigations.

In September 2011, this need was met by Cambridge Assessment's Research Division. As part of its wider research programme, a programme of studies to complement OCR's re-engagement work was set up, known as the HE Engagement research programme. This rolling programme was designed to extend over several years. From the start, it adopted systematic approaches to data collection and analysis, which were critical in giving credibility to the evidence upon which qualifications and curriculum development decisions could be based. Overall, a 'mixed methods' approach was used, to enable the corroboration and triangulation of findings drawn from both quantitative and qualitative data. In addition to generating an evidence base, the programme provides an example of how research in itself can be viewed as an important means of restoring and strengthening links between assessment organisations and HE.

The first phase of the research programme comprised three linked studies. They were designed to address four overarching questions which had surfaced in discussions within OCR's consultative forums:

1. In which areas do university lecturers think new undergraduates are most prepared?
2. In which areas do university lecturers think new undergraduates are least prepared?
3. What, therefore, are the transitional challenges for new undergraduates?
4. How could A levels be improved?

Study 1: What are the impacts of qualifications for 16 to 19 year olds on Higher Education? A survey of 633 university lecturers

Our first study (Suto, 2012; Suto, Mehta, Brown and Jeffery, 2012) comprised a survey of university lecturers in Biology, English, Mathematics, plus a diverse range of other subjects. The study's objectives were to collect mostly quantitative data indicating:

i. the perceived strengths and weaknesses of typical new undergraduates, and
ii. the systemic factors that are considered by lecturers to contribute to these skill sets and deficits.

We developed a questionnaire through an iterative piloting process involving lecturers and awarding body colleagues. The final version was available for completion on-line. It comprised 13 questions, took 10 minutes to complete, was suitable for lecturers in any subject, and did not refer to A levels from any particular awarding body.

Over 3000 lecturers were invited by e-mail to participate in the study, 633 of whom responded. Although all university groupings were targeted equally, 40% of respondents were teaching at Russell Group universities and 60% were teaching at universities in other groupings. Although this over-representation of the Russell Group was not ideal, we found throughout our data analysis that the views and experiences of Russell Group lecturers were broadly similar to those of other lecturers.

Several striking findings emerged from the study:

- There was a healthy appetite among lecturers for engagement in research exploring the transition from A level to HE.
- Over half of lecturers thought that new undergraduates are underprepared for degree level study.
- ICT, teamwork, presentation skills and intellectual curiosity were the skills and attributes most likely to be considered strengths of typical undergraduates when they begin degree level study.
- Most lecturers thought that academic writing, self-directed study, independent inquiry and research, and critical thinking skills are weaknesses of typical undergraduates when they begin degree level study.
- Depth of subject knowledge was also a concern for some lecturers.
- History, English and Mathematics were the A level subjects considered to provide the best preparation for degree level study by lecturers across a wide range of subjects.
- 60% of lecturers indicated that their institutions provide additional support classes for underprepared first-year undergraduates. Classes often focus on writing and independent learning skills.
- 72% of lecturers had had to adapt their teaching approaches to teach underprepared first-year undergraduates. This most frequently entailed covering more basic, fundamental or lower level content. Teaching higher level study skills, essay writing, and academic writing were common adaptations among English lecturers. Biology lecturers taught more numeracy and mathematical skills.
- 87% of lecturers thought that too much 'teaching to the test' was a major factor contributing to undergraduates being underprepared.
- Many changes to A level suggested by lecturers related to pedagogy and student learning, and included allowing less spoon-feeding and teaching to the test. Other suggestions included making examination questions less predictable and reducing re-sit opportunities.

Study 2: How effective are curricula for 16 to 19 year olds as preparation for university? A qualitative investigation of lecturers' views

In our second study (Mehta, Suto, and Brown, 2012) we collected detailed data through a series of seven events in four different regions of England. A total of 46 lecturers participated, from a range of disciplines and with varied teaching experience. Each session entailed two linked research activities. First, participants completed a written prioritisation task. They were presented with 10 core academic areas and were asked to indicate:

i. the two areas in which new undergraduates were least prepared; and
ii. the two areas in which new undergraduates were most prepared.

The responses were anonymised, analysed immediately, and reported back to the participants. This prioritisation data thereby provided stimulus material for the second activity: a focus group. The discussion schedule comprised open-ended questions with related prompts and probes.

A thematic analysis of the primarily qualitative data enabled several of the key findings from the questionnaire study to be corroborated. For example, critical and higher order thinking skills, academic writing skills, and independent inquiry and research skills were again identified as weaknesses of new undergraduates. The general view among the lecturers was that students joined university well prepared in examination techniques, but unable to carry out analytical tasks. Many lecturers believed that the grades achieved at A level did not really alter the picture. In their view, most of the new undergraduates, irrespective of previous grades, were not prepared for university study.

The participants emphasised that in addition to A level content, other aspects such as the structure of assessment and the opportunity for in-depth learning were equally important in preparing students for university study. A lack of preparedness was felt to result in a steep learning curve for the students, sometimes leading to students failing courses or dropping out of university. This challenge was considered to add to the financial, social, and personal challenges faced by new undergraduates. The participants described several of their solutions to transitional challenges. These included delivering extra classes for undergraduates (either stand-alone or integrated within wider courses) and using particular pedagogical techniques.

As in the questionnaire study, the participants perceived teaching at A level to involve a lot of spoon-feeding, to be narrow, and to focus on teaching to the test. They also thought that the modular structure of A levels and the many opportunities to re-sit examinations contributed to transitional challenges. (In 2000, a revised modular A-level structure was introduced, with twice-yearly opportunities to be assessed on each unit.) When reflecting on the research tasks, most participants felt that there should be more communication between schools and universities, since they realised that their knowledge about A levels was limited. They felt that such interaction would help them to increase their awareness of syllabuses and examinations, and ensure some common ground between HE and A level.

Study 3: A review of the literature examining the pedagogical differences between A level and university

The third study (Jeffery, 2012) comprised a review of recent research investigating pedagogical differences between A level and university from the perspectives of students, A level teachers, and university lecturers. Our aim was to identify research that explained or accounted for some of transitional challenges revealed in Studies 1 and 2. The review included research conducted in the UK from 2001 to 2010; that is, after the introduction of Curriculum 2000. Although only 10 studies met the full inclusion criteria of the review, all were comprehensive in their examination of the topic. Most studies employed self-report methods and the academic subjects covered included the English subjects, History, Geography, Psychology, and Business Studies.

A series of relevant differences between education at A level and at university emerged. For example, the main aims of A level pedagogy were widely considered by teachers and students to be to impart the curriculum and ensure that students obtain good grades. A level teachers saw their students as receptive learners, and in keeping with this view, they maintained personal and frequent interactions with them. At A level, essay-writing was seen to be about regurgitating the facts and figures needed to score highly in examinations. In contrast, at university the main pedagogical aim was perceived to be to encourage autonomy, self-confidence, problem-solving abilities and subject enthusiasm. Lecturers and tutors adopted a more 'hands-off' approach and saw themselves as responsible for providing only basic information, which students should investigate in greater depth for themselves. Essay-

writing was regarded as an opportunity to illustrate original and critical thinking.

These types of qualitative difference between the two educational levels are likely to be longstanding. However, potentially newer differences in approaches to assessment also surfaced. At A level, it was felt that students were thoroughly prepared for examinations and were provided with all the information they could possibly need. Assessment feedback was personalised and provided regularly and in a timely way. At university, on the other hand, students were expected to gather for themselves most of the assessment information they needed. Feedback was felt to be much more general and not always timely.

Policy developments in early 2012

The findings from the above three studies were initially disseminated at a national conference at the start of April 2012 (Suto *et al*., 2012a). This presentation coincided with the publication of a letter from the English Secretary of State for Education to the Chief Executive of the national qualifications regulator, Ofqual. The letter expressed concerns that although A levels have much to commend them, they fall short of commanding a desirable level of confidence within HE and among other stakeholders. The Secretary of State reiterated and expanded upon his earlier vision (Department for Education, 2010) for universities to be more involved in developing A levels:

> *I would therefore like to see universities having far greater involvement in the design and development of A level qualifications than they do at present. That involvement should be both when qualifications are developed initially and following each examinations cycle, so that universities' influence over the qualification develops over time… I am particularly keen that universities should be able to determine subject content, and that they should endorse specifications, including details of how the subject should be assessed.* (Department for Education, 2012a).

The Secretary of State went on to write:

> *The discussions I have had with university academics and school and college leaders on the subject of A-levels have left me concerned about the impact of the current modular structure on students' education, and their ability to make the connections between different topics within a subject that are so crucial for deep understanding. I will therefore be interested in your views – and in those of others – regarding A level modules (including the division of the qualification into AS and A2), and in particular the opportunity to take modules in January, together with the impact of resitting on confidence in A level standards*. (Department for Education, 2012a).

The Secretary of State's concerns were in line with those expressed by the lecturers in Studies 1 and 2. The letter and the research were reported jointly in much of the national media (for example, BBC, 2012). The Chief Executive of Ofqual replied to the letter, agreeing that the role of universities in the development of A level should increase. She clarified that:

> *Making sure that A levels are fit for purpose means getting four things right: subject content (curriculum), teaching, assessment, and level of demand* (Ofqual, 2012a).

The Chief Executive went on to agree with the Secretary of State that universities should be involved (together with awarding bodies and learned societies) in determining subject content, and to indicate that Ofqual would also take account of HE views on A level assessment. She also indicated that Ofqual would hold a public consultation on the structure and assessment arrangements for A level. The consultation was launched in June 2012, lasted three months, and sought views on the full range of proposed changes to the qualification, including:

i. the involvement of HE in the design and sign-off of A levels;
ii. the abolition of January exams and the limiting of re-sits; and
iii. whether or not AS levels should continue, with a range of options put forward.

In June 2012, following on from its earlier (2011) consultation paper, Cambridge Assessment produced its own policy paper (Cambridge Assessment, 2012), setting out how the Government could best support HE–awarding body interactions, and arguing that greater involvement of university academics in setting the content of A levels would both be a better guarantee of school standards and improve the university admissions process. The policy paper claimed that if university lecturers and tutors were to determine A level content, with awarding bodies focusing on how to test that knowledge, then the state's role in setting examination standards could be minimised. Arguably, such changes would end the constant tinkering with syllabuses that currently occurs, as examinations would only need to change when leading academics felt change was needed. The paper proposed that 'communities of interest' should be set up around each A level qualification. Such communities would bring together subject communities, learned societies, HE, schools and colleges, teachers, and awarding bodies to share a particular view of what constitutes the A level standard in relation to a subject. This would enable the standard of each qualification to be maintained over time by all those with a direct interest in it.

Study 4: Comparing the assessment types students encounter at A level and university

Whilst Ofqual's consultation was underway, a further phase of the HE Engagement research programme was launched. It comprised two studies which explored in greater depth some of the transitional issues raised in the first three studies. One study (Wilson, Child and Suto, 2013) began in July 2012. Its aim was to systematically compare the assessments that students encounter at A level and with those encountered in their first year at university. As in Study 1, this research focused on Biology, English and Mathematics. Assessment materials from 16 universities were collated, and we compared them to assessments in the equivalent A levels.

We were interested in four aspects of assessment at A level and university:

i. the types of assessment to which students were exposed (extended writing, multiple choice etc.)
ii. the written guidance and scaffolding students were given during the assessments
iii. the opportunity for re-sits
iv. the timings of the assessments throughout the academic year.

We found a greater variety of assessment types at university in comparison with at A level. Biology entailed the most varied assessment at university, followed by English and Mathematics. Interestingly, the

written guidance provided to students was greater and more detailed at university than at A level. It is possible that this guidance is in place due to the emphasis at university on working independently, and to help students negotiate new forms of assessment. Unlike at A level, university students were given only one re-sit opportunity in the majority of cases, with a cap on the potential mark that could be achieved. Students also had to cope with earlier summative assessment in comparison with A level.

Study 5: An exploration of additional support classes at university

In Studies 1 and 2, lecturers indicated that students arrived at university underprepared in certain skill areas including critical/higher order thinking, academic writing, and independent research. Furthermore, 60% of the lecturers in Study 1 reported that their institutions offered additional support to their students which focused on their general skills, subject-specific content or both. These findings gave rise to Study 5 (Mehta, Child, Brown, and Suto, 2013), which began in August 2012 and had three main aims:

i. to collect more in-depth information on the content and structure of additional support classes in Biology, English and Mathematics
ii. to gather the views of students, lecturers and A level teachers on the effectiveness of these classes
iii. to determine the potential for the content and skills covered in these classes to be included at A level.

For each target subject, case studies were conducted at three contrasting universities that ran additional support classes for their students. Each case comprised lecturer and student interviews, lesson observations by researchers and an A level teacher, and a facilitated discussion between the A level teacher and university lecturer.

We found that classes focused on a range of subject-specific and general skills. For example in Biology, classes typically centred around the appropriate writing and formatting of scientific reports, while in English, classes focused on developing students' awareness of critical approaches and independent reading. For Mathematics, a variety of specific content areas were covered, to fill perceived gaps in knowledge for students progressing from A level. The A level teachers thought that many of the areas covered in the additional support classes had been addressed at A level, but in less detail. In Biology, recommended changes to the A level included earlier development of project skills and numeric calculations. For English, participants advocated that the A level include a greater focus on independent criticism and historical background of texts. For Mathematics, participants suggested that making closer links between different content areas would be beneficial. More in-depth analysis of the data collected is currently underway.

November 2012: Ofqual announces the reduction of re-sit opportunities

In November 2012, whilst the above Studies 4 and 5 were underway, Ofqual published the outcomes of its consultation on A level reform. Almost a thousand stakeholders participated in the consultation, and key findings included support for:

- the principle of HE engagement with A level design, (however there was less support for universities specifically 'endorsing' each A-level)
- students being assessed at the end of each of their first and second year of study
- the removal of January examinations and reduced re-sit opportunities
- increasing synoptic assessment in A levels, allowing students to integrate and apply their skills, knowledge and understanding with breadth and depth
- reducing internal assessment.

(Ofqual, 2012b).

On the basis of these findings, Ofqual announced a significant interim reform: that from January 2014 onwards, there would be no January examination sessions for A levels. This meant that from September 2013, students would only be able to sit AS and A level examinations in the summer. Ofqual explained that this change was likely to be the first of wider reforms. It stated that:

> *Further changes to the system, including more involvement from higher education into the design of A levels, and changes to the A level structure, are also being considered and will be announced at a later date.* (Ofqual, 2012b).

Confirmation of A level reforms by the Department for Education

In late January 2013, these further changes were announced. The Secretary of State wrote another widely publicised letter to the Chief Executive of Ofqual, setting out his policy steers (DfE, 2013). According to this letter, new A level syllabuses will be developed which do not have the modular structure that was introduced in 2000. Instead there will be a return to linear assessment, with all examinations being taken at the end of the two-year course. The AS level examination will remain, however it will exist in a more traditional form. It will no longer be taken after the first year of A level study and will not count towards a full A level. Instead, AS levels will be exclusively stand-alone qualifications with half the content of full A levels. They are to be taken alongside A levels in other subjects after two years of study. The new A level and AS level syllabuses are to be introduced for teaching in schools from September 2015, and will initially be in all 'facilitating' subjects except languages; that is: English (Language, Literature and Language and Literature), Mathematics and Further Mathematics, Physics, Chemistry, Biology, History and Geography. The Secretary of State's letter also confirmed plans for 'leading' universities to be more closely involved in developing the content of new A levels, beginning with those subjects which are most commonly required for undergraduate study.

These changes to A level have received mixed responses. While the reversion to linear assessment fits with the preferences of many of the HE representatives who participated in our research programme, there is considerable concern over the reforms to AS levels. According to *The Daily Telegraph*, a spokesman for the University of Cambridge believed the loss of AS levels in their current form would damage the fairness of the admissions process, and in particular, the university's efforts to admit students from disadvantaged backgrounds. Without current AS levels, such students might lack the confidence to apply:

> *This change is unnecessary and, if implemented, will jeopardise over a decade's progress towards fairer access to the University of Cambridge…AS is the most reliable indicator available of an applicant's potential to thrive at Cambridge*. (*The Daily Telegraph*, 2013).

This view was shared by the Russell Group universities, whose Director General commented:

> *Whilst we have welcomed the Government's review of the modular structure of the A-level, we do not believe this need be extended to the complete removal of the AS examination from the A-level*. (Russell Group, 2013).

In March 2013, the Chief Executive of Ofqual responded to the Secretary of State's letter. She reported on an agreement between Ofqual and the awarding bodies to review a further six A level subjects in time for first teaching in 2015: Psychology, Art and Design, Sociology, Business Studies, Economics, and Computing. This means that overall, subjects covering approximately 70% of A level entries will be reviewed (Ofqual, 2012c). The regulator is also now exploring mechanisms for awarding bodies to work with HE to obtain agreement on A level content. At the time of writing, it is likely that some form of supervisory committee will be convened, possibly for each subject. In effect, this may 'officialise' some of the processes of stakeholder consultation and input gathering that Cambridge Assessment has developed and engaged in over the last few years.

Conclusion

To date, the HE Engagement research programme has dovetailed effectively with policy developments at a national level. It has been helpful in providing curriculum developers with an evidence base upon which to re-develop A levels in time for first teaching in September 2015. The research is likely to continue over the coming years, investigating new issues and questions as they become apparent. We believe that the use of systematic and methodologically sound approaches to research will strengthen the development of future qualifications.

References

Cambridge Assessment. (2012). *A better approach to Higher Education/Exam Board interactions for post-16 qualifications*. Available online at: http://www.cambridgeassessment.org.uk/ca/digitalAssets/193735_Cambridge_Assessment_HE_engagement_June_2011.pdf (Accessed 8 March 2013).

Department for Education. (2010). *The importance of teaching: Schools White Paper*. Available online at: https://www.education.gov.uk/publications/standard/publicationdetail/page1/CM%207980 (Accessed 8 March 2013).

Department for Education. (2012a). *Letter of 30th March 2012 from the Secretary of State for Education to the Chief Executive of Ofqual on the reform of GCE A levels*. Available online at: http://media.education.gov.uk/assets/files/pdf/s/secretary%20of%20state%20to%20ofqual%20on%20a%20levels.pdf (Accessed 8 March 2013).

Department for Education. (2012b) *Government publishes destination data for the first time. News webpage*. Available online at: http://www.education.gov.uk/inthenews/inthenews/a00211875/government-publishes-destination-data-for-the-first-time- (Accessed 8 March 2013).

Department for Education. (2013). *Letter of 22nd January 2013 from the Secretary of State for Education to the Chief Executive of Ofqual on the reform of GCE A levels*. Available online at: http://media.education.gov.uk/assets/files/pdf/l/ofqual%20letter%20alevels%20v2.pdf (Accessed 8 March 2013).

Jeffery, E. (2012). *A review of the literature examining the pedagogical differences between A level and university*. Cambridge Assessment report. April 2012.

Kingdon, M. (1991). *The Reform of Advanced Level*. London: Hodder & Stoughton.

Mehta, S., Suto, I., & Brown, S. (2012). *How effective are curricula for 16 to 19 year olds as a preparation for university? An investigation of lecturers' views*. Paper presented at the European Conference for Educational Research, University of Cádiz, Spain.

Mehta, S., Child, S., Brown, S. & Suto, I. (2013). *Thrown in at the deep end? Exploring students', lecturers' and teachers' views on additional support lessons at university*. Paper presented at the Future of Education Conference, Florence, Italy.

Ofqual. (2012a). *Letter of 3rd April 2012 from the Chief Executive of Ofqual to the Secretary of State for Education on the reform of GCE A levels*. Available online at: http://media.education.gov.uk/assets/files/pdf/l/letter%20to%20secretary%20of%20state%20re%20a%20level%20reform%203%20april%202012.pdf (Accessed 8 March 2013).

Ofqual. (2012b). *Ofqual announces changes to A levels*. Press release. Available online at: http://www.ofqual.gov.uk/news/ofqual-announces-changes-to-a-levels/ (Accessed 8 March 2013).

Ofqual. (2012c). *Letter of 21st March 2013 from the Chief Executive of Ofqual to the Secretary of State for Education on the reform of GCE A levels*. Available online at: http://www.ofqual.gov.uk/files/2013-03-22-letter-to-secretary-of-state-for-education-alevel-reform.pdf (Accessed 22 March 2013).

Raban, S. (2008). *Examining the world. A history of the University of Cambridge Local Examinations Syndicate*. Cambridge: Cambridge University Press.

Russell Group (2013). *AS level reform*. Press release. Available online at: http://www.russellgroup.ac.uk/russell-group-latest-news/154-2013/5450-aslevel-reform/ (Accessed 8 March 2013).

Suto, I. (2012). *How well prepared are new undergraduates for university study? An investigation of lecturers' perceptions and experiences*. Paper presented at the annual conference of the Society for Research in Higher Education, Newport, Wales.

Suto, I., Mehta, S., Brown, S. & Jeffery, E. (2012a). *How can the A level ensure a successful transition to university education? Emerging findings from related research with HE stakeholders.* Paper presented at the annual UCAS Admissions Conference, Birmingham, UK.

The Daily Telegraph. (2013). *Michael Gove's A-level reforms condemned by Cambridge*. Available online at: http://www.telegraph.co.uk/news/uknews/9822217/Michael-Goves-A-level-reforms-condemned-by-Cambridge.html (Accessed 8 March 2013).

Wilson, F., Child, S. & Suto, I. (2013). *A Comparison Of Assessment At School And University: More Than Just Increasing Demands*. Paper to be presented at the European conference for Educational Research, Istanbul, Turkey.

Comparing progression routes to post-16 Science qualifications

Carmen Vidal Rodeiro Research Division

Introduction

At present, awarding bodies in England provide schools and students with a wide choice of Science qualifications, aiming to ensure that pupils study Science that is relevant and up-to-date and that there is choice in the courses to prepare them for different routes post-16.

The most recent programme of study for Science at level 2[1] allows students to study a single GCSE[2] in Science. This single Science GCSE, which has great emphasis on scientific literacy, is intended to provide a good foundation to study the subject at AS or A level[3], and a good background for specialism in other sciences (Millar, 2006).

Additionally, pupils can take one of two complementary GCSEs: GCSE Additional Science or GCSE Additional Applied Science in order to cover a more comprehensive programme of study in Science. Additional Science, which has a more academic focus, prepares pupils further for progression to study AS and A levels in the Sciences. The Additional Applied Science has a strong focus on work-related learning.

Pupils can also study separately GCSE Biology, GCSE Chemistry and GCSE Physics to gain three full GCSEs in Science. In 2009 a change was introduced to the level 2 Science curriculum in England: all pupils who achieved over a certain threshold in Science tests at age 14 would be entitled to study the three separate Sciences at GCSE (Biology, Chemistry and Physics). The main motivation for establishing this entitlement was to get more students to study more Science at level 2, with the long term aim of increasing the supply of scientists, engineers and technologists in the workforce (Fairbrother and Dillon, 2009; HM Treasury *et al*., 2009).

Since 2010, the international GCSE, or IGCSE, has been accredited and funded in maintained schools in England and therefore maintained schools can offer this qualification in Biology, Chemistry and Physics to their students (independent schools in England had been offering IGCSEs in the Sciences prior to and in 2009). The IGCSE prepares students for further academic work, including progression to AS/A level study and to the International Baccalaureate programme.

There is also a vocational route in Science offered at level 2: Applied Science GCSE. This qualification, a double award, is designed to offer students the opportunity to widen their participation in vocationally related learning. The course is intended to provide students with the technical knowledge, skills and understanding needed in the workplace, in further education or training. In particular, it aims to provide students with the ability to apply their Science knowledge and skills to solving scientific problems in a variety of vocational contexts.

There are other Science qualifications at level 2, equivalent to one or more GCSEs, that account for a small percentage of the Science qualifications offered in schools (e.g. BTEC[4] First in Applied Science or Cambridge Nationals in Science). They are an alternative to the courses mentioned above and they intend to provide students with the technical knowledge and skills needed in the workplace, in further education or in training. Percentages of candidates taking them have been increasing over the last few years (Vidal Rodeiro, 2012b).

According to a report from the Office of Qualifications and Examinations Regulation (Ofqual, 2009), all Science subjects should give students opportunities to develop their interest in Science, develop a critical approach to scientific evidence and methods, acquire and apply skills, knowledge and understanding of 'how Science works' and its essential role in society and acquire scientific skills, knowledge and understanding necessary for progression to further learning. However, it has been argued recently (e.g. Wolf, 2011; Homer, Ryder and Donnelly, 2013) that some level 2 Science courses may not prepare students for the study of Science subjects at level 3. Consequently some students may decide not to pursue a Science subject post-16, or if they continue to study it they may drop the subject or not fulfil their potential in terms of the grades they achieve.

In recent years, there have also been some claims about Science options being restricted in some schools and the effect that this could have on students' futures (e.g. Banner *et al*., 2010). In fact, some young people have reported that they had compromised their choices by tailoring their options to what the schools could make available (e.g. Vidal Rodeiro, 2007; BBC News, 2009).

Due to the diversity in the level 2 Science curriculum, with a range of choices for students with different attitudes towards Science available, it is important to explore the different pathways taken by students who succeeded in the different options at level 2 in order to gain a better understanding about how level 2 qualifications are contributing to the supply of people with STEM[5] skills. In fact, a recent area of considerable interest in educational research has been the supply of young people gaining good A level grades in Science subjects, so that they would be eligible to continue on to university to study these subjects and to take up careers in Science or related fields.

Furthermore, in light of the Wolf review of vocational education (Wolf, 2011) evidence is needed to show if applied level 2 courses such as the BTEC First in Applied Science and the Cambridge National in Science provide meaningful pathways for studying Science at level 3 or whether

1. Each regulated qualification in England has a level between entry level and level 8. Qualifications at the same level are a similar level of demand or difficulty. To find out more about qualification levels visit http://www.ofqual.gov.uk/help-and-advice/comparing-qualifications/.
2. General Certificates of Secondary Education. These are the qualifications taken by the majority of 16 year olds in England.
3. AS and A levels are qualifications taken by students between the age of 16 to 18 in England. A levels are usually spaced out over two years and made up of two components: AS and A2 levels. AS levels can stand as a qualification on their own or can be carried on to A2 to complete a full A level qualification.
4. BTEC stands for *Business and Technology Education Council*, which used to award the qualification. BTECs are now awarded by the Edexcel exam board.
5. Science, Technology, Engineering and Mathematics.

students taking more academic pathways are better prepared for further study.

A recent report by the Royal Society (The Royal Society, 2011) recommends: 1) that the increasing diversity of Science qualifications needs to be reviewed and its impact on the numbers of students taking Science post-16 evaluated; and 2) that awarding organisations should make available detailed data on the participation, attainment and progression of students taking their specifications in Science.

This research carries out an in-depth statistical analysis of examination data from level 2 and level 3 Science qualifications, focusing on participation, attainment and progression. The central questions to be addressed in this article are:

- What is the provision of level 2 Science qualifications in English schools?
- What are the characteristics of the candidates taking level 2 Science qualifications?
- To which level 3 Science qualifications do students with level 2 qualifications in Science progress?
- What is the performance at level 3 of candidates progressing from different level 2 Science qualifications? Do some progression routes produce better outcomes than others?

Data and methods

Data

To answer these research questions, data from the National Pupil Database on uptake and performance at levels 2 and 3 for the examination years 2008/09 and 2010/11 were used.

The National Pupil Database (NPD), compiled by the Department for Education, is a longitudinal database for all children in schools in England, linking student characteristics to school and college learning aims and attainment. The NPD holds pupil and school characteristics such as age, gender, ethnicity, level of deprivation, attendance and exclusions, matched to pupil level attainment data (Key Stage 2 to Key Stage 5 assessments and other external examinations). Students who start in a school/college are only recorded on the NPD if they enter for a qualification; those who leave school/college after a short time or do not sit examinations are not present in the data.

Methods

The analyses presented in this article were carried out in two stages:

- Stage 1 consisted of a comprehensive analysis of the provision of level 2 Science in English schools and the characteristics of the candidates taking each level 2 Science qualification;
- Stage 2 investigated the progression from level 2 Science qualifications to level 3 study.

STAGE 1

Provision of level 2 Science subjects

The research addressed this issue through descriptive analyses that looked into schools' characteristics such as their type and their level of deprivation.

School type information was obtained from the NPD and schools were classified according to the following types:

- comprehensive schools
- academies
- independent schools
- selective schools
- secondary modern schools
- other.

A deprivation score for each school was calculated as the average score on the Income Deprivation Affecting Children Index (IDACI)[6] of the students attending it. This score, available in the NPD, shows the percentage of children in the Lower Super Output Area (LSOA)[7] in which the student resides who live in families that are income deprived.

The schools' deprivation variable was continuous and therefore was divided into three equal-sized groups (low, medium, high) using percentile values, and schools were classified accordingly.

Characteristics of the candidates taking Science qualifications at level 2

The research addressed this issue through descriptive analyses that focussed on:

- prior attainment of candidates;
- socio-economic background of candidates (level of deprivation);
- type of centre where the level 2 qualifications were obtained.

The prior attainment of candidates taking Science qualifications at level 2 was measured by the total marks obtained in the Science Key Stage 3 tests taken at age 14. Key Stage 3 refers to the three years of schooling when pupils are aged between 11 and 14. All pupils in this Key Stage must follow a programme of education in at least 15 areas. At the end of this stage, pupils are tested and are awarded attainment levels depending on what they are able to do. These tests cover English, Mathematics, Science and ICT[8].

The socio-economic background of the students was measured by the Income Deprivation Affecting Children Index (IDACI), available in the NPD (see above for details about this index).

Students were classified into six different groups depending on the type of school in which they obtained the Science qualification at level 2 (as above).

STAGE 2

The second stage of the research looked into the progression from level 2 Science qualifications to level 3 study. This involved a follow-up of candidates who completed level 2 Science qualifications in 2009 and a series of descriptive analysis looking into their progression routes to level 3.

It should be noted that there is a limitation regarding the data used for this research. The linking between candidates with Science qualifications in 2009, and candidates recorded in the Key Stage 5 extract of the 2011 NPD, was carried out using a unique pupil number common to both

6. See page 19 of http://www.communities.gov.uk/documents/communities/pdf/733520.pdf for a detailed explanation of this index.
7. A LSOA is a conglomeration of a number of census output areas (each output area has about 150 households). They usually have a minimum population size of 1000 and an average of 1500. There are over 34000 LSOAs in England.
8. It should be noted that until 2008, performance at Key Stage 3 was assessed by a series of externally-marked tests. However, from 2009, results from the Key Stage 3 tests have no longer been available for the entire cohort of students in maintained schools in England.

datasets. However, some candidates did not have this unique identifier and therefore matching in this way was not possible for them. As a result, some candidates who progressed to qualifications at level 3 might not have been included in the analyses.

More detailed analyses on progression to the traditional A level qualifications in Biology, Chemistry and Physics (which make up the majority of Science level 3 qualifications) were also carried out and are reported in Vidal Rodeiro (2012a). In particular, uptake and performance of the three individual A level subjects were investigated for different groups of students using background student and school variables. Among candidates' features, Vidal Rodeiro (2012a) looked at gender, overall attainment at level 2, level of deprivation, uptake of GCSE Mathematics alongside Science qualifications and attainment in Science at level 2. In terms of schools, it looked at attributes such as the type, the attainment, the composition with respect to gender and whether the school had a sixth form or not.

Results

Provision of level 2 Science qualifications

In this article, provision of a subject or qualification was defined as the number or percentage of schools with at least one student entering for examination in the subject or qualification. It could be argued that this may not cover all of the provision as schools might offer a subject, but none of their students study it or take an examination in it. However, there was no data available that would allow capturing the definite subject/qualification offer in schools and only provision as defined above was considered.

Table 1 presents the provision of level 2 qualifications/subjects in Science offered for certification in the academic year 2008/09 in secondary schools in England. The table shows that GCSEs in Biology, Chemistry and Physics were offered in about 40% of the secondary schools. The provision figures for the GCSE in Science and the GCSE in Additional Science (about 80% and 70%, respectively) were higher than for the GCSEs in Biology, Chemistry and Physics. The provision figures for the Applied Science subjects were much lower than for the academic subjects. For example, the Additional Applied Science GCSE was offered in around 17% of the schools and the BTEC Firsts and the OCR Nationals[9] in Applied Science were offered, respectively, in around 11% and 3% of schools.

Table 1: Level 2 Science related qualifications/subjects, June 2009

Science qualification/ subject at level 2	*Number of schools*	*Percentage of schools*
GCSE Science	4331	80.4
GCSE Additional Science	3711	68.9
GCSE Additional Applied Science	935	17.4
Applied Science GCSE	382	7.1
GCSE Biology	2301	42.7
GCSE Chemistry	2182	40.5
GCSE Physics	2177	40.4
GCSE Biology: Human	101	1.9
GCSE Environmental Science	132	2.5
GCSE Geology	45	0.8
OCR National Applied Science	144	2.7
BTEC First Applied Science	574	10.7

9. Now called Cambridge Nationals.

Table 2 presents the percentages of schools offering Science subjects by the type of school. The provision of the three separate Sciences (Biology, Chemistry, and Physics) was higher in selective schools than in any other type of school. It should be noted that there might be some independent schools that offered the IGCSE in the separate Sciences rather than the GCSE for certification in 2009, despite IGCSEs not being accredited or funded for state maintained schools until 2010. Table 2 also shows that provision for the Science and Additional Science GCSEs was lower in independent and selective schools than in comprehensive schools, academies and secondary modern schools.

The percentages of selective and independent schools offering a BTEC First in Applied Science were very small (below 1%), compared to the percentages of comprehensive schools, academies or secondary modern schools doing so (around 20%). The provision figures for the OCR National in Applied Science were lower than the figures for the BTEC but the patterns of provision by school type were similar (about 5% of comprehensive schools and academies and fewer than 1% of selective and independent schools offered the qualification).

Table 2: Level 2 Science related qualifications/subjects, June 2009 – by type of school

Science qualification/ subject at level 2	*Compre-hensive*	*Academy*	*Indepen-dent*	*Selective*	*Secondary Modern*	*Other*
GCSE Science	98.6	95.0	73.0	85.6	96.8	48.5
GCSE Additional Science	97.3	93.3	67.1	81.1	94.2	12.7
GCSE Additional Applied Science	31.3	22.3	2.6	6.3	28.6	2.3
Applied Science GCSE	12.4	11.5	0.8	0.0	9.7	1.2
GCSE Biology	54.3	56.4	55.0	82.9	36.4	9.4
GCSE Chemistry	52.6	55.7	53.5	82.9	27.3	6.0
GCSE Physics	53.1	55.1	52.8	82.9	26.6	5.6
GCSE Biology: Human	2.1	1.4	0.3	0.9	3.2	2.6
GCSE Environmental Science	3.9	3.2	1.0	0.0	3.9	0.8
GCSE Geology	1.1	1.4	0.6	2.7	0.0	0.1
OCR National Applied Science	4.5	5.1	0.6	0.0	3.2	0.3
BTEC First Applied Science	17.6	20.0	0.9	0.0	18.8	1.9

There was a strong pattern in the provision of Science by school deprivation (see Table 3). Higher percentages of schools with pupils living in areas of low deprivation provided GCSEs in the three separate Sciences than schools with pupils living in areas of high deprivation (53% compared with 30%). Furthermore, the percentages of schools offering OCR Nationals and BTECs increased with increasing deprivation. This might lead to restricted options for high ability students who live or attend schools in deprived areas.

Table 3: Level 2 Science related qualifications/subjects, June 2009 – by school level of deprivation

Science qualification/ subject at level 2	*Low deprivation*	*Medium deprivation*	*High deprivation*
GCSE Science	87.5	89.2	85.0
GCSE Additional Science	76.9	79.0	72.8
GCSE Additional Applied Science	22.7	23.7	21.2
Applied Science GCSE	6.7	10.3	10.3
GCSE Biology	55.2	44.2	31.9
GCSE Chemistry	52.9	42.3	29.9
GCSE Physics	52.8	42.2	30.1
GCSE Biology: Human	2.0	2.6	1.1
GCSE Environmental Science	4.2	3.7	1.2
GCSE Geology	1.8	1.0	0.1
OCR National Applied Science	2.1	3.2	5.0
BTEC First Applied Science	7.6	12.9	20.9

Characteristics of candidates taking level 2 Science qualifications

In this section of the article the characteristics of the candidates taking Science qualifications at level 2 are investigated. For this purpose, students taking Science qualifications at level 2 were classified as pursuing the following routes:

- GCSE Science
- GCSE Science & GCSE Additional Science
- GCSE Science & GCSE Additional Applied Science
- GCSE Science & GCSE Additional Science & GCSE Additional Applied Science
- GCSE Biology & GCSE Chemistry & GCSE Physics
- Applied Science GCSE
- BTEC First Applied Science
- OCR National Applied Science

Prior attainment

Figure 1 shows the mark distribution in the Key Stage 3 Science tests for candidates taking different Science qualifications at level 2 and reveals that there were differences in the Science prior attainment among the different level 2 Science routes, with the highest average prior attainment corresponding to those students taking the separate Sciences (Biology, Chemistry, Physics) and the lowest prior attainment to those taking Science only or the BTEC and OCR National routes. An analysis of variance showed that the differences between the prior attainment of candidates following the different Science routes were statistically significant. Figure 1 also shows that the variation in the Key Stage 3 scores was smaller among the students taking the separate Sciences than among the students following any other route.

These results therefore suggest that prior attainment in Science may have an effect on the type of Science qualification pursued at level 2.

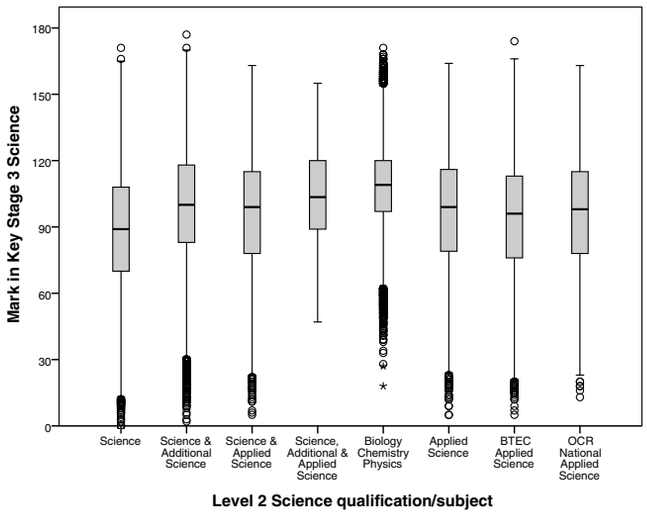


Figure 1: Prior attainment (Key Stage 3 Science) – by Science route at level 2

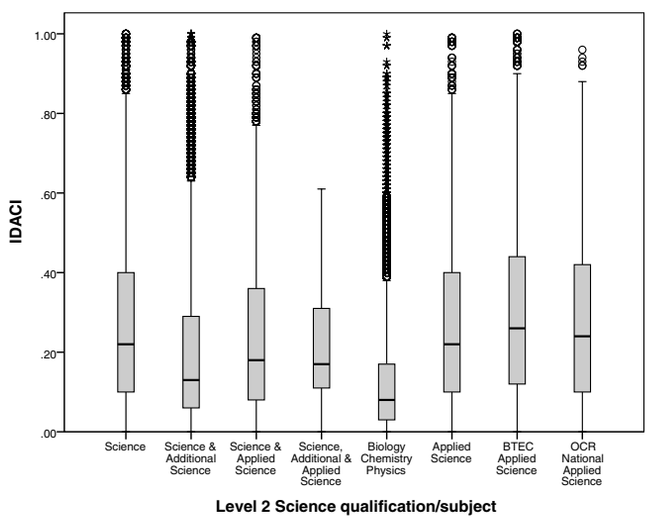


Figure 2: Level of deprivation (IDACI) – by Science route at level 2

Level of deprivation

Figure 2 shows the level of deprivation (IDACI) for candidates taking different Science qualifications at level 2 and reveals that there were differences in the level of deprivation among the different level 2 Science routes, with the lowest level of deprivation corresponding to those students taking the separate Sciences and the highest level of deprivation to those taking the BTEC and OCR National routes. An analysis of variance showed that the differences between the level of deprivation of candidates following the different Science routes at level 2 were statistically significant. As above, it is worth noting that the variation in the deprivation scores was lower among the students taking the separate Sciences than among those following any other route.

Type of school

Table 4 shows the type of school where the level 2 Science qualification was obtained (the table shows column percentages). For example, in independent schools, 35% of the candidates had pursued the triple Science route. This contrasts with only 11% in comprehensive schools. Similarly, only 1% of candidates in independent schools obtained a BTEC in Applied Science, whilst 6% of the candidates in comprehensive schools and 7% of the candidates in academies did so.

Table 4: Type of school where the level 2 Science route was pursued (column %)

Science route at level 2	*Comprehensive*	*Academy*	*Independent*	*Selective*	*Secondary Modern*	*Other*
Science	12.8	12.2	7.0	1.7	19.9	64.1
Science & Additional Science	51.6	45.2	40.3	36.7	43.9	12.9
Science & Additional Applied Science	8.3	5.8	1.3	0.5	7.9	1.2
Science, Additional & Additional Applied Science	0.0	0.0	0.0	0.1	0.0	0.3
Biology & Chemistry & Physics	10.5	15.0	34.7	50.7	4.3	7.6
Applied Science GCSE	3.5	3.3	0.3	0.0	3.4	1.5
BTEC First Applied Science	5.7	7.4	1.2	0.0	5.8	1.4
OCR National Applied Science	0.8	1.3	0.2	0.0	0.3	0.7

Progression to post-16 qualifications in Science

Uptake of level 3 qualifications and level 3 Science qualifications

This section of the article investigates the progression from candidates with at least one Science qualification at level 2, obtained in June 2009, to level 3 qualifications certificated in June 2011.

Figure 3 shows that 49% of candidates at the end of Key Stage 4[10] in 2009 obtained a level 3 qualification at the end of June 2011. This percentage was slightly higher, 53%, among candidates with at least one Science qualification at level 2.

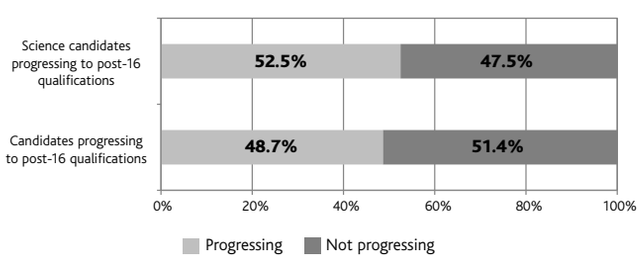


Figure 3: Progression from level 2 to post-16 (level 3) qualifications

Table 5 shows the percentages of candidates progressing to any level 3 qualification by the Science route at level 2 and Table 6 the percentages progressing to level 3 Science qualifications. These tables highlight that the route with higher progression rates to level 3 was the triple Science route (Biology, Chemistry and Physics), with 80% of the candidates who pursued it progressing to level 3 and 46% of the candidates progressing to a Science qualification at level 3. Table 6 highlights that around 26% of the candidates taking Science and Additional Science GCSEs progressed

10. Key Stage 4 refers to the two years of schooling in England when pupils are aged 15 to 16.

Table 5: Candidates progressing to level 3 qualifications (any subject) – by Science route at level 2

Science route at level 2	*Progressing to level 3 (any subject)*	
	Number	*%*
Science	17255	22.5
Science & Additional Science	167507	58.7
Science & Additional Applied Science	14795	35.3
Science, Additional & Additional Applied Science	85	50.3
Biology & Chemistry & Physics	63637	80.0
Applied Science GCSE	6843	36.9
BTEC First Applied Science	11251	33.6
OCR National Applied Science	1832	38.1

Table 6: Candidates progressing to Science qualifications at level 3 – by Science route at level 2

Science route at level 2	*Progressing to level 3 (Science)*	
	Number	*%*
Science	775	1.0
Science & Additional Science	37287	13.1
Science & Additional Applied Science	959	2.3
Science, Additional & Additional Applied Science	22	13.0
Biology & Chemistry & Physics	36467	45.8
Applied Science GCSE	747	4.0
BTEC First Applied Science	1107	3.3
OCR National Applied Science	169	3.5

to level 3 Science qualifications and that fewer than 5% of the candidates following an applied route at level 2 (Applied Science GCSE, Science and Additional Applied Science, BTEC First or OCR National) progressed to Science at level 3.

Table 7 shows the percentages of candidates[11], by Science route at level 2, who progressed to specific qualifications at level 3. The key findings from this table are:

- very similar percentages of candidates progressed to A level and BTEC qualifications from GCSE Science;
- higher percentages of candidates from level 2 applied routes (including from Applied Science GCSEs, OCR Nationals and, particularly, from BTEC Firsts) progressed to level 3 BTEC qualifications than to any other qualification;
- much higher percentages of candidates from the triple Science (Biology, Chemistry, Physics) or the double Science (Science and Additional Science) routes progressed to A level than to any other qualification;

11. Row percentages do not necessarily add up to 100% because candidates can progress to more than one qualification at level 3.

Table 7: Candidates progressing to level 3 qualifications (in any subject area) – by Science route at level 2

Science route at level 2	*Level 3 qualifications*															
	A level	*AS level*	*Applied A level*	*Applied AS level*	*Applied AS/A level*	*Pre-U*	*EPQ*	*Diploma*	*IB*	*AQA Bacc*	*NVQ*	*VRQ*	*BTEC*	*OCR National*	*Key Skills*	*Other*
Science	44.3	28.3	10.1	4.9	0.1	0.0	3.4	0.7	0.2	0.2	0.6	7.5	44.0	3.2	1.8	0.1
Science & Additional Science	69.5	39.4	8.6	3.6	0.1	0.1	5.0	0.5	0.5	0.7	0.3	4.4	24.2	1.7	1.9	0.0
Science & Additional Applied Science	43.7	27.9	13.3	6.4	0.1	0.0	3.3	0.7	0.1	0.1	0.3	7.9	45.3	3.2	2.1	0.1
Science, Additional & Additional Applied Science	64.7	37.6	17.6	2.4	0.0	0.0	3.5	1.2	0.0	1.2	0.0	1.2	35.3	0.0	1.2	0.0
Biology & Chemistry & Physics	88.1	51.8	3.7	1.6	0.0	0.5	7.1	0.1	1.5	1.7	0.1	1.7	7.0	0.6	1.6	0.0
Applied Science GCSE	44.6	27.6	16.4	6.7	0.1	0.0	4.6	0.6	0.1	0.1	0.4	7.4	45.7	4.9	1.5	0.0
BTEC First Applied Science	33.5	22.8	9.1	3.8	0.1	0.0	2.9	0.7	0.3	0.1	0.7	8.5	60.2	3.4	2.4	0.0
OCR National Applied Science	40.0	27.3	10.8	6.0	0.3	0.0	3.4	1.2	2.3	0.1	0.3	6.8	47.8	5.3	2.5	0.1

Table 8: Candidates progressing to level 3 Science qualifications – by Science route at level 2 (as a percentage of the candidates taking the subject at level 2 and progressing to level 3)

(a) Progression to AS/A level in Science subjects

Science route at level 2	*Level 3 Science qualifications*									
	AS[12] level Biology	*AS[12] level Chemistry*	*AS[12] level Physics*	*AS level Environmental Science*	*A level Biology*	*A level Chemistry*	*A level Physics*	*A level Environmental Science*	*AS level Applied Science*	*A level Applied Science*
Science	0.8	0.5	0.4	0.0	1.1	0.5	0.5	0.1	0.2	0.3
Science & Additional Science	2.5	1.6	1.1	0.2	10.9	7.1	4.8	0.4	0.4	0.7
Science & Additional Applied Science	0.6	0.3	0.1	0.1	0.9	0.3	0.2	0.1	0.9	2.3
Science, Additional & Additional Applied Science	2.4	2.4	2.4	25.9	15.3	10.6	3.5	1.2	0.0	2.4
Biology & Chemistry & Physics	4.2	3.8	2.9	0.2	31.5	27.9	18.6	0.5	0.2	0.3
Applied Science GCSE	0.7	0.4	0.2	0.1	1.4	0.6	0.4	0.2	1.4	5.0
BTEC Applied Science	0.8	0.4	0.2	0.0	0.7	0.5	0.3	0.0	0.4	0.8
OCR National Applied Science	0.3	0.2	0.1	0.1	1.1	0.6	0.5	0.1	1.4	2.5

(b) Progression to other level 3 qualifications in Science

Science route at level 2	*Level 3 Science qualifications*						
	Pre-U[13] Biology	*Pre-U Chemistry*	*Pre-U Physics*	*IB Biology*	*IB Chemistry*	*IB Physics*	*BTEC Applied Science*
Science	0.0	0.0	0.0	0.1	0.0	0.0	0.8
Science & Additional Science	0.0	0.0	0.0	0.3	0.2	0.1	1.0
Science & Additional Applied Science	0.0	0.0	0.0	0.0	0.0	0.0	1.8
Science, Additional & Additional Applied Science	0.0	0.0	0.0	0.0	0.0	0.0	3.5
Biology & Chemistry & Physics	0.0	0.0	0.0	0.9	0.8	0.4	0.6
Applied Science GCSE	0.0	0.0	0.0	0.1	0.0	0.0	2.0
BTEC Applied Science	0.0	0.0	0.0	0.1	0.0	0.0	6.6
OCR National Applied Science	0.0	0.0	0.0	1.2	0.3	0.1	2.8

12. Candidates progressing to AS level and not continuing to A level.
13. Note that percentages have been rounded to 0.0 so there might be some candidates who progressed to Pre-U Science subjects, particularly from the triple Science route at level 2.

- higher percentages of candidates from the Applied Science routes than from the more traditional routes progressed to applied AS/A levels;
- higher percentages of candidates from the triple Science route than from almost any other route progressed to the International Baccalaureate (IB) Diploma (1.5% compared to fewer than 1%);
- higher percentages of candidates from the triple Science route progressed to the Extended Project Qualification (EPQ) than from any other route (around 7% compared to less than 5%);
- very small percentages of level 2 Science candidates progressed to level 3 OCR Nationals. Indeed, the highest percentage was around 5% and corresponded to the progression of those candidates who obtained an OCR National at level 2.

Table 8 shows the percentages of candidates, by Science route at level 2, who progressed to specific Science qualifications at level 3. The key findings from this table are:

- higher percentages of candidates from the triple Science or the double Science routes progressed to A level in Biology, Chemistry or Physics than to any other qualification;
- independently of the level 2 Science route, progression was generally higher to A level in Biology than to A level in Chemistry or Physics;
- there was hardly any progression to A levels in Biology, Chemistry or Physics from the level 2 applied routes;
- higher percentages of candidates from the Applied Science routes than from the more academic routes progressed to the level 3 BTEC in Applied Science.

Performance in level 3 Science qualifications

This section of the article reports on the performance in Science at level 3 of candidates progressing from each of the level 2 Science routes.

Performance in each of the different Science routes at level 3 has been measured in a different scale. For example, A levels are graded A* to E whilst BTECs are graded as pass, merit or distinction. Therefore, the figures presented in this section do not allow comparison across the level 3 qualifications and comparisons should only be made within qualifications.

Figures 4 to 6 display the average performance in the more popular level 3 Science qualifications by Science route at level 2. Figure 4 highlights that in each level 3 academic qualification (A levels in Biology, Chemistry and Physics) those progressing from triple Science at level 2 did better than anybody else. However, when interpreting these results, it should be taken into account that candidates from different level 2 routes might have different levels of academic ability, measured in this work by the Key Stage 3 tests results. For example, candidates taking the triple Science route might be more academically able than those taking the single Science route and therefore, their performance at level 3 is likely to be better. This issue has been taken into account in Vidal Rodeiro (2012a) where uptake and performance of level 3 Science qualifications were investigated for different groups of students.

14. Grades were converted into points using the UCAS tariff. Available online at: (http://www.ucas.com/students/ucas_tariff/tarifftables/)

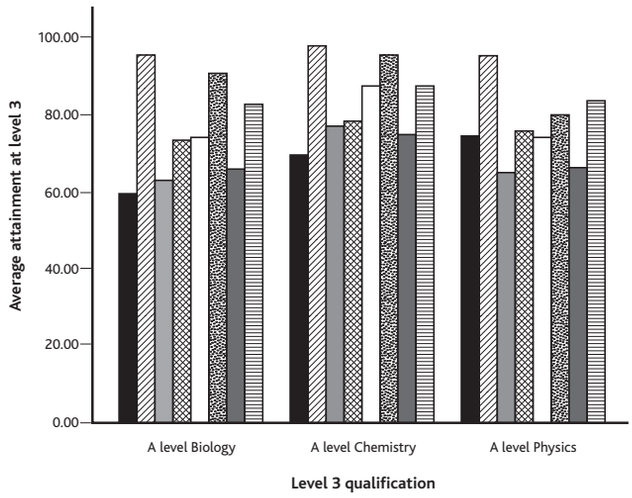


Figure 4: Average performance in A level Biology, Chemistry and Physics[14]

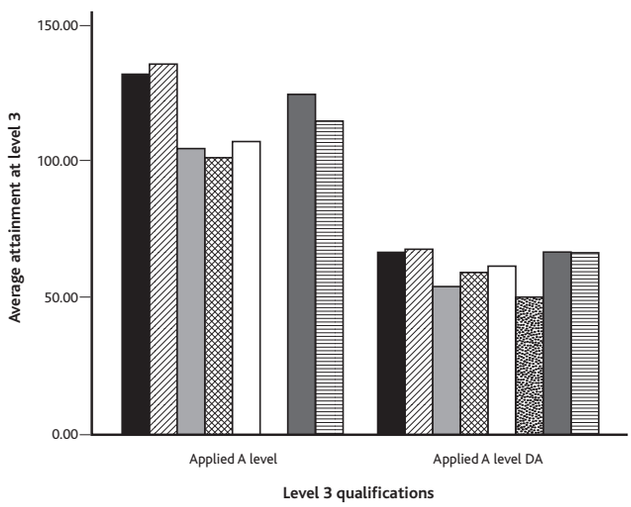


Figure 5: Average performance in applied A level and applied A level double awards in Science[14]

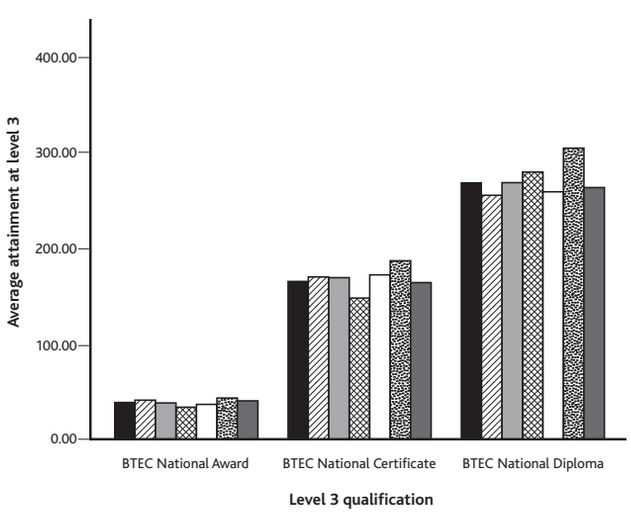


Figure 6: Average performance in BTEC qualifications in Applied Science[14]

Common key to Figures 4, 5 and 6

- Applied Science GCSE
- Biology – Chemistry – Physics
- BTEC Applied Science
- OCR National Applied Science
- Science
- Science, Additional & Additional Applied Science
- Science & Aditional Applied Science
- Science & Additional Science

Figures 5 and 6 present the performance in some of the applied level 3 Science qualifications. In particular, Figure 5 highlights that performance in the applied A levels was better among those students progressing from the triple Science, but followed closely by the performance of those progressing from the Applied Science and the Science and Additional Applied Science routes at GCSE. Figure 6 shows that performance on BTEC qualifications at level 3 was a little better among students progressing from the Science and Additional Applied Science GCSE than among students progressing from any other level 2 route.

The above results show that the Additional Applied Science, which has a strong focus on work-related learning, was associated with better results in applied level 3 options than other qualifications at level 2.

Conclusions and discussion

This research aimed to gather detailed information about pupils taking level 2 Science qualifications and their progression to level 3 on completion of specific Science routes. Its main driver came from national concerns about the relatively small numbers of young people choosing to study Science subjects beyond compulsory schooling.

The outcomes of this research showed that the different level 2 Science qualifications enabled learners to progress to level 3 in a variety of ways.

The key conclusions of the analyses presented in this article are summarised and discussed briefly below.

Provision of level 2 Science qualifications

Although Science is compulsory at Key Stage 4, there were alternative routes available at level 2 in schools and colleges in England to fulfil the Science curriculum requirements.

The provision figures provided in this article showed that the majority of the secondary schools in England offered Science and Additional Science GCSEs. However, there has been a shift in recent years towards separate GCSEs in Biology, Chemistry and Physics (see, for example, Vidal Rodeiro (2012b) for an account of the provision in Science qualifications at level 2 from 2009 to 2011). This shift might have been due to the Government's commitment in February 2009 to increase access to triple Science in order to ensure that all pupils in maintained schools had access to triple Science GCSEs by 2014 and that the percentage of state school pupils taking the three Science subjects reached 17% by 2014.

There were also other alternative level 2 Science qualifications, such as the Applied Science GCSE (double award) or the BTEC First and OCR National in Applied Science which were offered in a small percentage of secondary schools in England.

This research showed that schools with certain characteristics were associated with a higher provision of particular Science options at level 2. In particular, provision of GCSEs in Biology, Chemistry or Physics was higher in selective schools or in schools with favourable economic factors (e.g. their students had low IDACI scores, indicating low levels of deprivation). Furthermore, provision of Applied Science qualifications, such as the Applied Science GCSE, the OCR National in Applied Science or the BTEC First in Applied Science, was greater in comprehensive schools and academies and in schools where pupil deprivation was high.

It should be noted that the fact that not all Science qualifications (or subjects) were on offer in the majority of schools may restrict students' options. For example, high ability students who live or attend a school in a deprived area would have to tailor their choices to what the school is offering. Furthermore, it should be borne in mind that there is no information about how the selection of students into the different Science routes takes place within the schools. It might be that there is no actual choice made by the students and it is possible that, particularly at level 2, schools determine which students are suitable for each route.

Characteristics of candidates taking level 2 Science qualifications

This research showed clear differences in the background of the students taking the different Science routes at level 2.

First, the outcomes of this research revealed that prior attainment in Science (at Key Stage 3) might have had an effect on the type of Science qualification pursued at level 2. In particular, the research showed that the highest prior attainment corresponded to students taking separate GCSEs in Biology, Chemistry and Physics and the lowest prior attainment to students taking GCSE Science, BTEC First in Applied Science or OCR National in Applied Science.

Secondly, there were differences in the level of deprivation between level 2 Science routes, with the lowest level of deprivation corresponding to students taking GCSEs in Biology, Chemistry and Physics and the highest level of deprivation corresponding to those taking a BTEC First or an OCR National in Applied Science.

Finally, given the numbers of schools of each type in England, and their provision of Science qualifications/subjects, it is not surprising that the overwhelming majority of the candidates obtained level 2 Science qualifications in comprehensive schools. However, the research revealed differences in the proportions of candidates in each type of school who followed each level 2 Science route. For example, in independent schools over a third of the candidates pursued the triple Science route (GCSEs in Biology, Chemistry and Physics); this contrasts with only 10% of the candidates in comprehensive schools and 50% in selective schools. Similarly, only 1% of candidates in independent schools pursued a BTEC First in Applied Science, whilst 6% of the candidates in comprehensive schools did so.

Progression to post-16 qualifications in Science

One of the main aims of this research was to investigate the progression from candidates with at least one Science qualification at level 2 (obtained in June 2009) to level 3 qualifications, certificating in June 2011. Through analysis of national data this research showed that the level 2 Science route with the highest progression rates to level 3 was the triple Science route, with around 46% of the candidates progressing to a Science qualification at level 3. Only around 26% of the candidates from the double Science route progressed to level 3 Science qualifications and fewer than 5% of the candidates following an applied route (applied GCSE, Science and Additional Applied Science, BTEC First in Applied Science or OCR National in Applied Science) progressed to Science at level 3.

Performance in Science subjects at level 3 was better for pupils progressing from separate GCSEs in Biology, Chemistry and Physics than for pupils progressing from any other level 2 Science qualification. It should be noted though that pupils progressing from the triple Science route were more able (in terms of their Key Stage 3 Science results) than those progressing from other routes.

The above statements might suggest that choices at level 2 determine post-16 trajectories and therefore further study or employment options.

In particular, applied routes in Science did not show much progression to level 3 'academic' qualification/subjects and therefore progression to higher education could be restricted for candidates following those routes.

References

Banner, I., Donnelly, J., Homer, M.S. & Ryder, J. (2010). The impact of recent reforms in the key stage 4 science curriculum. *School Science Review*, **92**, 339,101–109.

BBC News (2009). *School rapped over GCSE sciences*. Available online at: http://news.bbc.co.uk/1/hi/education/8080139.stm (Accessed 15 January 2013).

Fairbrother, R. & Dillon, J. (2009). Triple science back on agenda. *School Science Review*, **91**, 334, 65–69.

HM Treasury, Department for Innovation, Universities and Skills and Department for Children, Schools and Families (2009). *Annual innovation report 2008*. London: HM Treasury.

Homer, M.S., Ryder, J. & Donnelly, J. (2013). Sources of differential participation rates in school science: the impact of curriculum reform. *British Educational Research Journal*. Available online at: http://onlinelibrary.wiley.com/doi/10.1080/01411926.2011.635783/full(Accessed online on 15 January 2013).

Millar, R. (2006). Twenty First Century Science. Insights from the design and implementation of a scientific literacy approach in school science. *International Journal of Science Education*, **28**, 13, 1499–1521.

Ofqual (2009). *The new GCSE science examinations: Findings from the monitoring of the new GCSE science specifications, 2007 to 2008*. Coventry: Office for Qualifications and Examinations Regulation.

The Royal Society (2011). *Preparing for the transfer from school and college science and mathematics education to UK STEM higher education: A 'state of the nation' report*. London: The Royal Society.

Vidal Rodeiro, C.L. (2007). *A level subject choice in England: patterns of uptake and factors affecting subject preferences*. Research Report. Cambridge: Cambridge Assessment.

Vidal Rodeiro, C.L. (2012a). *Comparing progression routes in science*. Research Report. Cambridge: Cambridge Assessment.

Vidal Rodeiro, C.L. (2012b). *Provision of science at level 2, 2011*. Statistics Report Series 39. Cambridge: Cambridge Assessment.

Wolf, A. (2011). *Review of Vocational Education – The Wolf Report*. London: Department for Education.

Early entry GCSE candidates: Do they perform to their potential?

Tim Gill Research Division

Introduction

There has been concern recently that a large and increasing number of candidates are certificating for GCSEs at a younger age than scheduled. Although there has been a trend of increasing early entry[1] over recent years (Gill, 2010), there has been a particularly large increase in the last two years (Department for Education, 2011). It is thought that the driver for this increase is the scrapping of the Key Stage 3 (KS3) tests (the final tests were taken in 2008). These were taken in year 9, and their absence means schools can now start teaching some GCSE subjects in this year, and subsequently enter candidates at the end of year 10, or in the winter sessions of year 11. One possible reason for early certification is that candidates can 'bank' a grade in the subject (in particular a grade C would contribute towards achieving the league table target of 5 A* to C grades as well as other important school accountability measures) allowing more time in year 11 to concentrate on other subjects. The concern is that many of these candidates are not reaching their potential in the subject because they certificate before they are ready. For example, Vidal Rodeiro & Nadas (2012) found evidence that candidates taking a (modular) English GCSE early were less likely to achieve a high grade than those certificating at the normal time.

1. This article refers both to 'early entry' and 'early certification'. However, they both mean the same thing – namely certificating for a GCSE before the summer of year 11.

A further issue relates to participation and performance in the subject at A level. Candidates who certificate early (particularly if they do so in June of year 10 or earlier) usually then have a break in studying the subject, meaning that they may lose interest or feel less confident that they are prepared for further study. Those students who do go on to take the A level in the subject may struggle because of this break.

This research first explores the extent of early certification in GCSEs. Two different years of data are used, to see whether there have been any changes in early certification patterns over time. Concerns about candidates not achieving their potential are then investigated by looking at GCSE performance based on certification session, as well as participation and performance in the same (or similar) subject at A level. Specifically, the research questions were:

- **RQ1** What are the patterns of early certification in GCSE subjects in recent years?
- **RQ2** How do candidates who certificate in a GCSE prior to June of year 11 perform relative to candidates taking the GCSE in June of year 11, after accounting for prior attainment?
- **RQ3** What effect does certificating in a GCSE early have on uptake and performance in the same (or a related) subject at A level, after accounting for prior attainment?

Data

The data for the analyses were taken from the National Pupil Database (NPD). This is a database of candidate level attainment and personal characteristics compiled by the Department for Education from data supplied by centres and awarding bodies. Two different extracts were used; the Key Stage 4 (KS4) extract which records all attainment by candidates at the end of KS4, and the Key Stage 5 (KS5) extract which records all attainment at A level or equivalent by candidates aged 16 to18.

For the GCSE analysis (research questions 1 and 2) KS4 extracts from 2009 and 2011 were used, to enable comparisons over time to be made. Each extract includes all candidates at the end of KS4 (i.e. in year 11) who took at least one qualification equivalent to a GCSE. All results for these candidates are recorded, including those taken in prior years or sessions. For example, if a candidate who was at the end of KS4 in 2011 took a GCSE in 2010 (i.e. when they were in year 10) this result would be recorded in the 2011 KS4 extract along with their results of GCSEs taken in 2011. If the candidate then re-sat the GCSE they took a year early in 2011 the result would also be recorded in the 2011 extract.
A variable for the year and session (summer or winter) in which the candidate certificated in each qualification is also included, enabling early certification in a subject to be identified as well as whether or not a candidate re-sat a GCSE.

It should be emphasised that the cohorts here are defined by age, not by exam year. Thus, when discussing the number of candidates in the 2011 cohort entering early for a GCSE this is not the number who entered early in 2011, but refers to the number of candidates who were in year 11 (i.e. aged 16) in 2011 and entered early for a GCSE (i.e. in a previous year).

For the A level analysis (research question 3) candidates in the 2009 KS4 extract were followed up at A level (in 2011). In order to do this the KS4 extract from 2009 was matched to the KS5 extract from 2011 using the unique candidate identifier in the extracts. This made it possible to investigate the uptake and performance at A level of candidates taking GCSEs in different sessions or years.

Results

Entry patterns

Table 1 presents the number of entries for all GCSEs in different years and sessions for the two different cohorts. It should be noted that the winter session is referred to here as January (which is the month in which OCR has its main winter session), but this also includes different winter sessions for the other boards (November and March for AQA and EdExcel for example). Sessions prior to June of year 10 were combined as very few candidates certificate before year 10.

Table 1: GCSE entries by year and session (all subjects)

Session	*2009 (n)*	*2009 (%)*	*2011 (n)*	*2011 (%)*
June Y11	4,486,279	91.8	4,108,947	85.0
Jan Y11	137,602	2.8	385,827	8.0
June Y10	230,834	4.7	290,391	6.0
Pre-June Y10	30,369	0.6	51,343	1.1
All	**4,885,084**		**4,836,508**	

This shows that the percentage of early entries increased substantially between 2009 (8.2%) and 2011 (15.0%). The largest increase was in entries in the winter of year 11, which went up from 2.8% in 2009 to 8.0% in 2011.

Tables 2 and 3 present the most common subjects to be taken early by candidates in the two different cohorts. For these tables, any re-sits have been removed, so only the first sitting of the exam is included. The tables present all entries in the subject for that cohort of students, the number of early entries and the number of early entries prior to year 11. This is to distinguish between subjects where most of the early entries were in January of year 11 and subjects where most of the early entries were prior to year 11.

The top ten subjects were almost identical in both years, with English and Maths having the highest numbers of early entries. This was particularly the case in 2011, where 41.1% of all entries in English and 36.9% of all entries in Maths were early. It is also interesting to note that in 2011 there were very high numbers of entries in these subjects in the same year (i.e. winter session 2011). Core Science and Statistics were the next most popular subjects to be taken early, but all the early entries in Statistics and almost all in Core Science were before year 11. Indeed for all the other subjects in the tables, almost all early entries were pre-year 11. Thus it is only English and Maths that are taken in the winter of year 11 in large numbers.

Table 2: Ten most common early entry subjects (2009 cohort)

Subject	*Total entries (n)*	*Early entries (n)*	*Early entries pre Y11 (n)*	*Early entries (%)*	*Early entries pre Y11 (%)*
English	600,084	107,131	25,425	17.9	4.2
Mathematics	599,738	62,910	30,993	10.5	5.2
Science (Core)	466,277	56,674	51,359	12.2	11.0
Statistics	72,503	36,441	36,441	50.3	50.3
French	168,557	16,579	16,579	9.8	9.8
English Literature	484,092	13,997	13,012	2.9	2.7
Religious Studies	169,366	10,243	10,243	6.0	6.0
Media/Film/TV Studies	60,928	4,660	4,660	7.6	7.6
German	68,862	4,553	4,553	6.6	6.6
Office Technology	33,355	4,182	4,182	12.5	12.5

Table 3: Ten most common early entry subjects (2011 cohort)

Subject	*Total entries (n)*	*Early entries (n)*	*Early entries pre Y11 (n)*	*Early entries (%)*	*Early entries pre Y11 (%)*
English	591,740	243,094	65,578	41.1	11.1
Mathematics	612,083	225,885	77,082	36.9	12.6
Science (Core)	359,331	37,088	35,901	10.3	10.0
Statistics	61,600	29,943	29,943	48.6	48.6
French	142,624	16,506	16,506	11.6	11.6
English Literature	449,778	14,543	14,156	3.2	3.1
Religious Studies	196,849	12,706	12,706	6.5	6.5
Media/Film/TV Studies	53,355	5,566	5,566	10.4	10.4
Spanish	59,048	5,088	5,088	8.6	8.6
German	58,594	4,884	4,884	8.3	8.3

School type

It is interesting to consider whether candidates in different types of school are more likely to certificate early. Almost all candidates attended one of five different types of school at the end of KS4; comprehensive, academy, independent, grammar or secondary modern. Tables 4 to 7 present the percentage of candidates within each of these school types entering English or Maths GCSE in different years and sessions (first certification only).

In both subjects and both years, candidates in academy, comprehensive and secondary modern schools were more likely to certificate early than those in independent or grammar schools. This difference was particularly stark in the January year 11 session. The differences may be partly due to the influence of school league tables and other accountability measures, with non-independent schools hoping to get candidates to 'bank' a grade C in a subject early so as to concentrate on other subjects. In all school types there were increases in the percentages of candidates certificating early in 2011 compared to 2009, but these increases were larger in academy, comprehensive and secondary modern schools.

Table 4: Percentage of candidates in each school type first certificating in each session (English 2009)

Exam session	*School type*				
	Academy	*Compre-hensive*	*Independent*	*Secondary Modern*	*Grammar*
June Y11	81.9	80.8	95.2	76.3	94.0
Jan Y11	13.5	15.2	2.8	17.8	3.0
June Y10	4.6	3.9	1.9	5.8	3.0
Pre-June Y10	0.0	0.1	0.1	0.1	0.0

Table 5: Percentage of candidates in each school type first certificating in each session (English 2011)

Exam session	*School type*				
	Academy	*Compre-hensive*	*Independent*	*Secondary Modern*	*Grammar*
June Y11	55.4	55.7	88.9	49.9	84.0
Jan Y11	31.0	33.2	7.6	33.8	9.8
June Y10	12.9	10.8	3.2	15.0	6.0
Pre-June Y10	0.8	0.3	0.4	1.2	0.2

Table 6: Percentage of candidates in each school type first certificating in each session (Maths 2009)

Exam session	*School type*				
	Academy	*Compre-hensive*	*Independent*	*Secondary Modern*	*Grammar*
June Y11	88.3	89.7	94.0	87.0	92.6
Jan Y11	6.4	5.5	1.6	6.9	0.8
June Y10	5.1	4.4	4.1	5.3	6.6
Pre-June Y10	0.3	0.3	0.2	0.7	0.0

Table 7: Percentage of candidates in each school type first certificating in each session (Maths 2011)

Exam session	*School type*				
	Academy	*Compre-hensive*	*Independent*	*Secondary Modern*	*Grammar*
June Y11	57.8	62.0	83.0	59.3	84.0
Jan Y11	26.4	26.1	9.9	21.4	5.5
June Y10	13.9	10.3	6.5	15.8	10.2
Pre-June Y10	1.9	1.5	0.8	3.2	0.4

Grades and progression

Research questions 2 and 3 involved exploring the relationship between the session(s) in which a GCSE is certificated and;

i. the performance in the subject;
ii. the likelihood of taking an A level in the subject and;
iii. the performance at A level in the subject or a related subject.

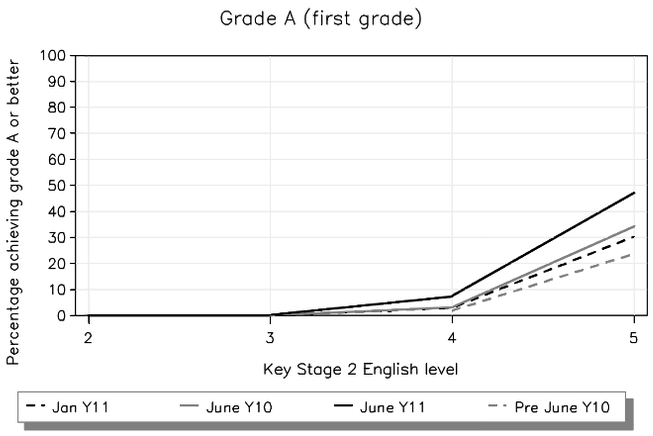
GCSE performance

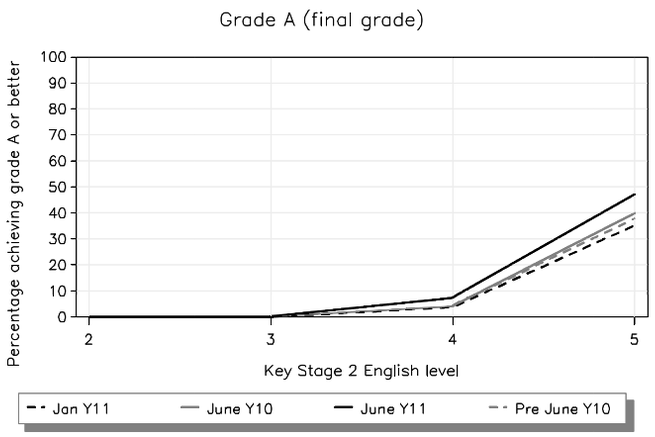
To explore the performance at GCSE it was decided to focus on the two most popular early certification subjects; English and Maths. Two separate analyses were undertaken, looking at the performance of candidates the first time they certificated in the subject and the final time they certificated, by first exam session. This means it is possible to observe the impact of certificating early, whilst also taking into account the effect of re-sitting. Performance was summarised by the percentage of candidates achieving at least a grade A in the subject and the percentage of candidates achieving at least a grade C.

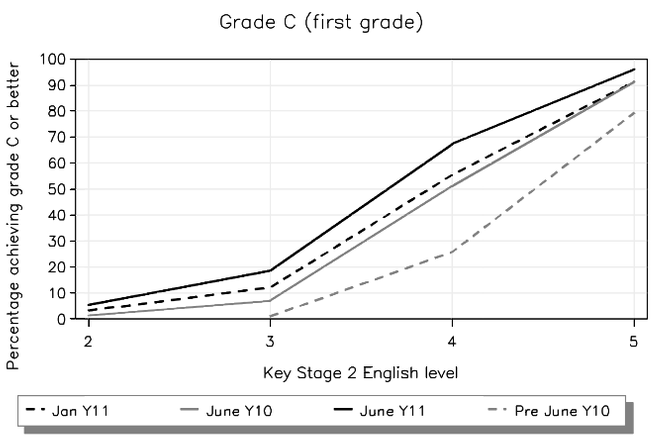
The prior attainment measure used for this analysis was the Key Stage 2 (KS2) test level achieved by the student in the relevant subject. There are issues with using this measure: first, the tests were taken five years prior to the scheduled date for GCSEs; and secondly, pupils in independent schools are not required to take these tests, meaning that there is no prior attainment data for many pupils in these schools. For example, in the 2011 cohort 61.8% of pupils in independent schools had KS2 results in English, compared with 95.9% of pupils in comprehensive schools. However, since no other prior attainment data is available, achievement in KS2 tests should at least give some indication of pupils' ability levels. KS2 levels range from 2 to 5, with pupils 'expected' to reach level 4 by the end of KS2.

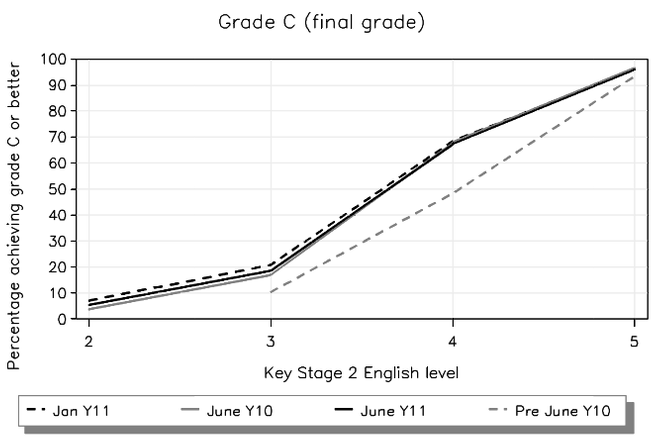
Figure 1 presents the results for English GCSE for the 2009 cohort and Figure 2 for the 2011 cohort. In each figure the graphs on the left hand side present the percentage of candidates achieving each grade or better the first time they certificated (by first exam session and KS2 level). The graphs on the right hand side show the percentage of candidates achieving each grade or better as their final (or best) grade (by first exam session and KS2 level). The table presents the same data. To give an example, in the 2009 cohort (Figure 1) 34.5% of candidates who took English for the first time in June of year 10 achieved a grade A or better in that exam session. However, 40.0% of these candidates achieved a grade A or better as their final grade (i.e. taking re-sits into account).

The first graph in both Figure 1 and Figure 2 show that candidates in both cohorts who certificated early were clearly less likely to achieve a grade A or better in that certification than those (with the same KS2 level) who certificated for the first time in June of year 11. Looking at the graphs for final grade, candidates certificating for the first time in



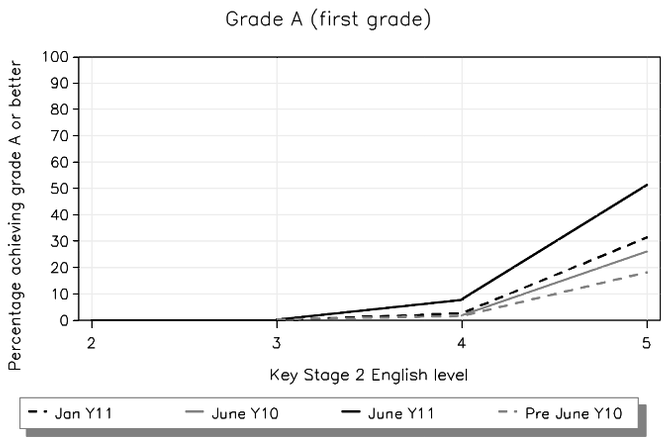


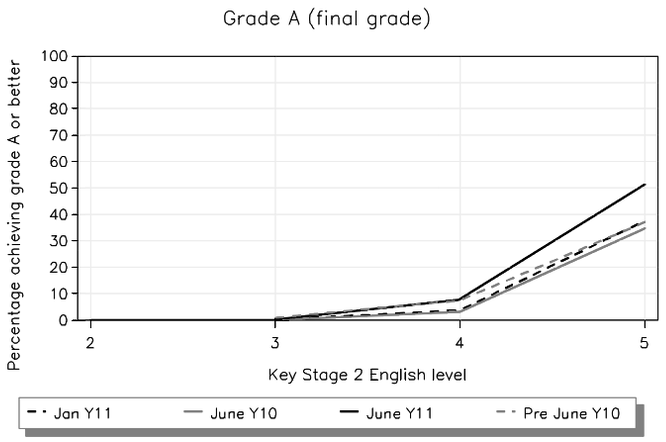


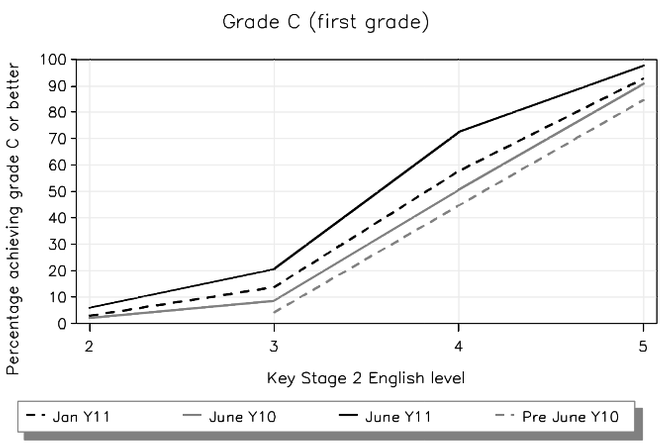


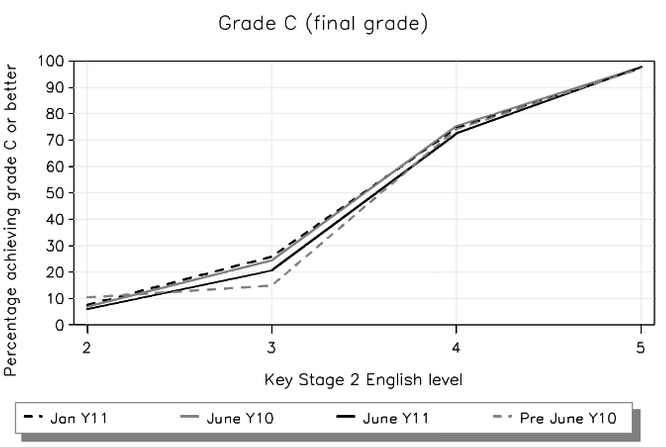
		Category	*KS2 English level*			
			2	*3*	*4*	*5*
Grade A	First grade	Pre-June Y10	0.0	0.0	1.9	23.7
		June Y10	0.0	0.1	3.2	34.5
		Jan Y11	0.0	<0.1	3.0	30.4
		June Y11	0.1	0.2	7.5	47.1
	Final grade	Pre-June Y10	0.0	0.0	4.5	37.9
		June Y10	0.0	0.1	4.2	40.0
		Jan Y11	0.0	0.1	3.8	35.2
		June Y11	0.1	0.2	7.5	47.1
Grade C	First grade	Pre-June Y10	0.0	1.0	25.8	79.3
		June Y10	1.4	7.1	51.1	91.3
		Jan Y11	3.3	12.2	55.5	91.3
		June Y11	5.4	18.7	67.4	96.1
	Final grade	Pre-June Y10	0.0	10.4	48.4	93.4
		June Y10	3.7	17.0	67.9	96.5
		Jan Y11	7.1	20.9	68.5	96.1
		June Y11	5.4	18.7	67.4	96.1

Figure 1: English 2009 – Percentage achieving key grades, by KS2 level and first certification session









		Category	*KS2 English level*			
			2	*3*	*4*	*5*
Grade A	First grade	Pre-June Y10	0.0	0.4	1.7	18.2
		June Y10	0.0	0.0	1.8	26.0
		Jan Y11	0.0	0.1	2.7	31.5
		June Y11	0.1	0.2	7.9	51.4
	Final grade	Pre-June Y10	0.0	0.8	7.5	37.1
		June Y10	0.0	0.1	3.2	34.8
		Jan Y11	0.0	0.1	3.7	37.6
		June Y11	0.1	0.2	7.9	51.4
Grade C	First grade	Pre-June Y10	0.0	4.2	44.7	84.7
		June Y10	2.2	8.7	50.8	90.9
		Jan Y11	2.9	13.8	57.8	92.9
		June Y11	6.0	20.7	72.6	97.7
	Final grade	Pre-June Y10	10.5	14.9	74.2	97.3
		June Y10	6.9	24.5	75.4	97.7
		Jan Y11	7.5	25.9	74.6	97.5
		June Y11	6.0	20.7	72.6	97.7

Figure 2: English 2011 – Percentage achieving key grades, by KS2 level and first certification session

June of year 11 were still most likely to get a grade A or better, although the differences were smaller.

Similarly, those certificating early were less likely to achieve a grade C or better in their first certification than those (with the same KS2 level) certificating for the first time in June of year 11. However, when looking at the final grade achieved a different picture emerges. For the 2009 cohort, there was almost no difference in the percentages achieving grade C or better for each (first) certification session. Only those who certificated prior to June of year 10 were less likely to get a grade C or better in their final grade. For the 2011 cohort there was barely any difference in the percentages amongst candidates with level 4 or 5 at KS2 for the different first certification sessions. However, for those with level 3 only at KS2, the candidates who certificated for the first time in June of year 10 or January of year 11 were most likely to get a grade C or better.

Figures 3 and 4 present the results for Maths for the 2009 and 2011 cohorts respectively. Looking at the results of first certification only for the 2009 cohort, candidates certificating in June of year 10 or June of year 11 were the most likely to achieve a grade A or better, particularly amongst those with level 5 in their KS2 Maths. When looking at final grade however, those who first certificated prior to June of year 10 or in June of year 10 were more likely to get a grade A than those certificating for the first time in June of year 11. For the 2011 cohort, those certificating for the first time in June of year 11 were clearly the most likely to get a grade A or better when considering first certification grade only. This was also the case when considering final grade, although those certificating first in June of year 10 were almost as likely. Those certificating first prior to June of year 10 or in January of year 11 were less likely to get a grade A or better in their final grade, particularly amongst those with level 5 at KS2.

Turning to the percentages getting at least a grade C in both the 2009 and 2011 cohorts, those certificating first in June of year 11 performed best when considering first certification grade only (at all levels of KS2). However, when looking at final grade this is not the case. For candidates in the 2009 cohort with level 3 at KS2, the highest percentage getting a grade C or better was amongst those certificating first in January of year 11. For those with level 4 at KS2, candidates certificating first prior to June of year 10 were most likely to get a at least a grade C. For candidates with level 5 at KS2 there was barely any difference due to the first certification session. In the 2011 cohort the differences were very small, at each level of KS2.

In summary, there is evidence that candidates certificating early for GCSEs in English and Maths tended to perform worse in their first certification than those certificating at the expected time. This is not very surprising as a lot of these candidates will not have had time to study the subject in enough depth to perform to their potential. For the best performing candidates this pattern was evident in terms of final grade as well, with the percentage achieving at least a grade A lower for candidates certificating early (with the exception of the 2009 cohort in Maths).

A level uptake and performance

In order to investigate the uptake and performance at A level of candidates certificating early for a GCSE, it was necessary to merge the KS4 NPD extract from 2009 with the KS5 extract from 2011. This should capture the majority of candidates in the 2009 cohort who went on to take an A level in the relevant subject.

The same two GCSE subjects (English and Maths) were used for this analysis. However, at A level there are three separate English subjects; English Language & Literature, English Language and English Literature. The most popular of these, by some distance, was English Literature, so it was decided to look at this subject separately, as well as investigating uptake and performance in all of the subjects combined (i.e. 'any English A level').

As well as looking at performance in A level Maths, it was also proposed that performance in A level Sciences might be affected by when Maths GCSE was taken. Since the Sciences (particularly Physics and Chemistry) rely heavily on mathematical knowledge, it may be that students who took Maths early would struggle in their Science A levels. Therefore, candidates taking A level Sciences were matched back to their GCSE results to determine when their Maths GCSE was taken. This analysis was undertaken for performance at A level only, not uptake.

The matched datasets were reduced further, for two reasons. Candidates getting below a grade C in the relevant GCSE were excluded as there were very few candidates who went on to take an A level in the subject. Secondly, there were very few candidates who took English GCSE prior to June of year 10, and even fewer of these went on to take any A levels. Therefore, candidates in this category were excluded.

A level uptake

Table 8 presents the number of matched candidates who took a GCSE in the listed subject and went on to take at least one A level or equivalent, and the percentage of these going on to take an A level in the subject (or a related one).

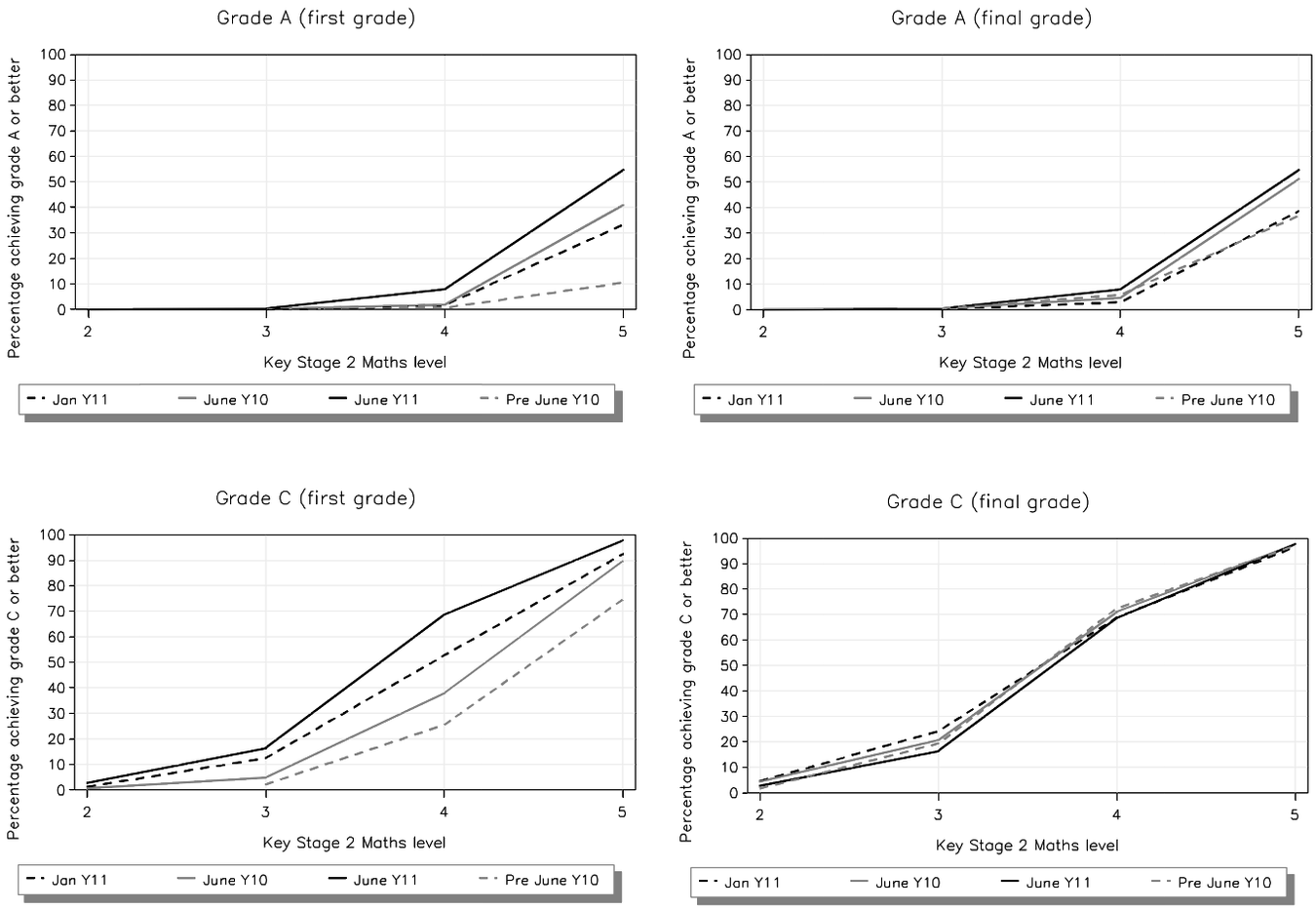
Table 8: Number of matched candidates and percentage taking A level

GCSE Subject	*A level subject*	*Candidates taking at least one A level*	*Candidates taking A level in subject (%)*
English	Any English	228,997	30.8
English	English Literature	228,997	17.2
Maths	Maths	218,400	24.8

Separate analyses were undertaken classifying candidates by when they first certificated in the subject at GCSE and when their final certification was. This was because both may impact on decisions about further study in the subject: candidates taking the GCSE early may feel that their study was rushed and therefore they did not enjoy it or do not feel it prepared them for further study, and this perception may persist even if they re-sit. Similarly, those whose *final* certification session was 'early' will have a gap before A level and thus may lose interest in the subject or not feel prepared for further study. Candidates were also classified by their final GCSE grade in the subject, as those who perform better at GCSE are more likely to go on to take the subject at A level.

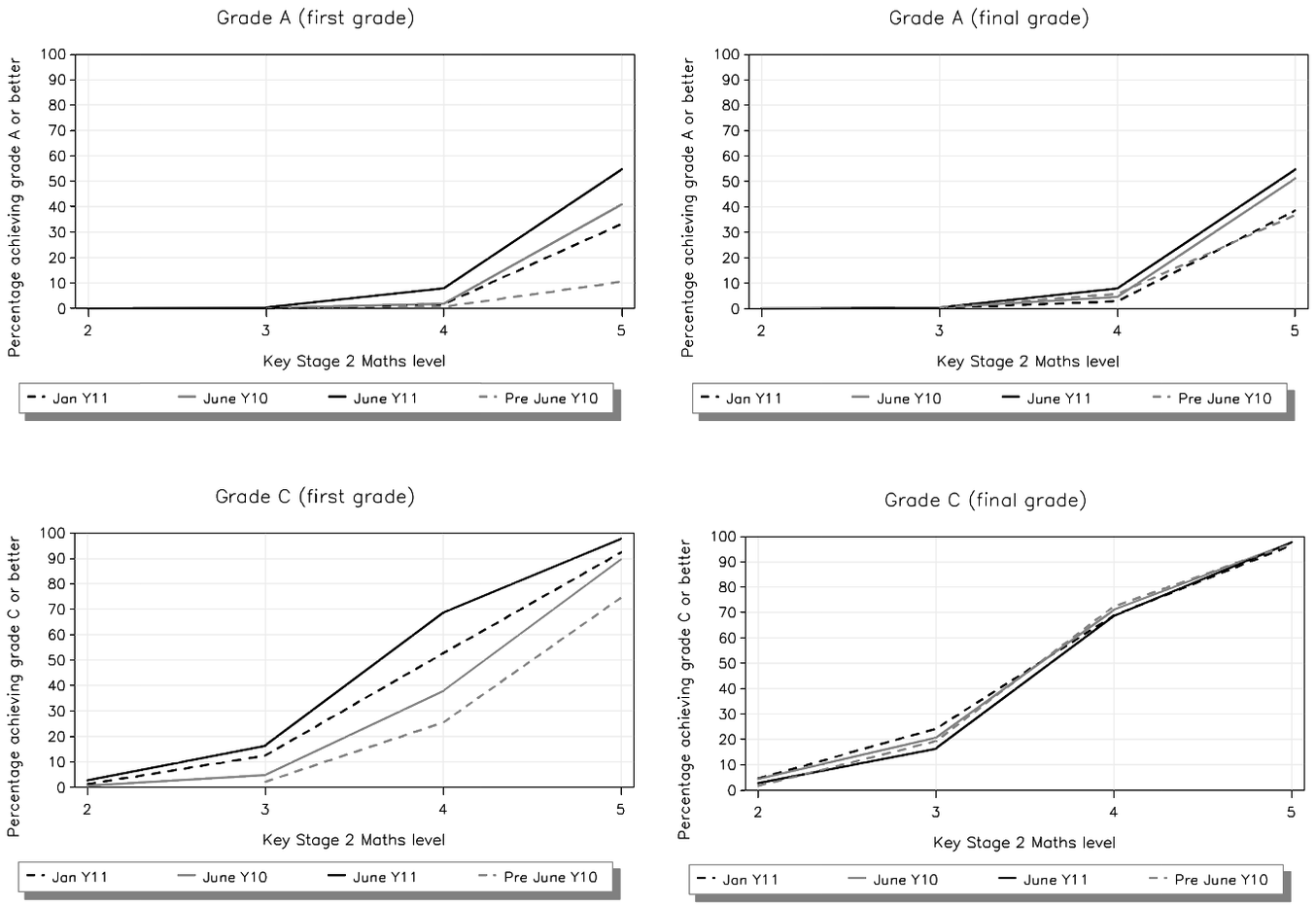
Figure 5 presents the percentage of candidates achieving each of grades A* to C at GCSE going on to take any English A level. Figure 6 presents the percentages going on to take English Literature A level. In each figure the left-hand graph is for the *first* certification session for the GCSE and the right-hand graph for the final certification session.

Figure 5 shows that candidates certificating in English GCSE for the first time in January of year 11 or June of year 10 were more likely to go on to take an English A level than those (with the same GCSE grade) who certificated for the first time in June of year 11. This pattern holds when looking at final certification session. Figure 6 shows that those certificating in June of year 10 were more likely to go on to take an A level in English Literature (although the differences were very small).



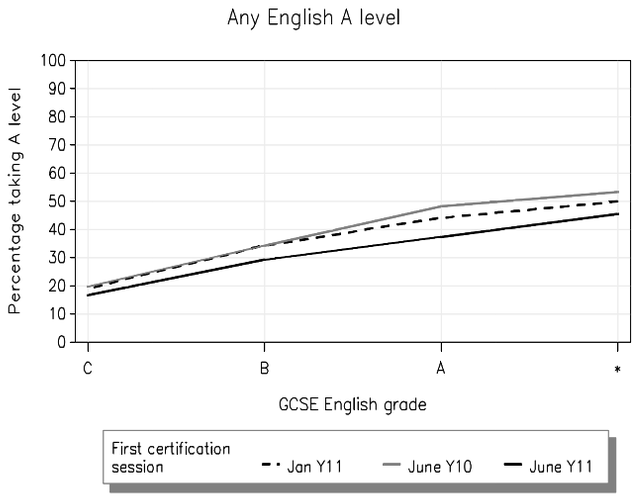
		First certification session	*KS2 Maths level*			
			2	*3*	*4*	*5*
Grade A	First grade	Pre-June Y10	0.0	0.0	2.4	20.6
		June Y10	0.0	<0.1	3.4	49.4
		Jan Y11	0.4	0.1	1.3	28.5
		June Y11	<0.1	0.2	5.2	45.8
	Final grade	Pre-June Y10	0.0	0.0	13.2	48.2
		June Y10	0.0	0.1	5.2	57.2
		Jan Y11	0.4	0.2	1.8	32.8
		June Y11	<0.1	0.2	5.2	45.8
Grade C	First grade	Pre-June Y10	0.0	4.4	49.4	82.9
		June Y10	0.0	3.5	40.3	92.4
		Jan Y11	0.8	10.3	48.5	87.6
		June Y11	1.6	13.4	62.6	96.0
	Final grade	Pre-June Y10	0.0	8.8	75.2	97.2
		June Y10	0.3	9.3	59.5	97.3
		Jan Y11	0.8	19.6	65.1	94.3
		June Y11	1.6	13.4	62.6	96.0

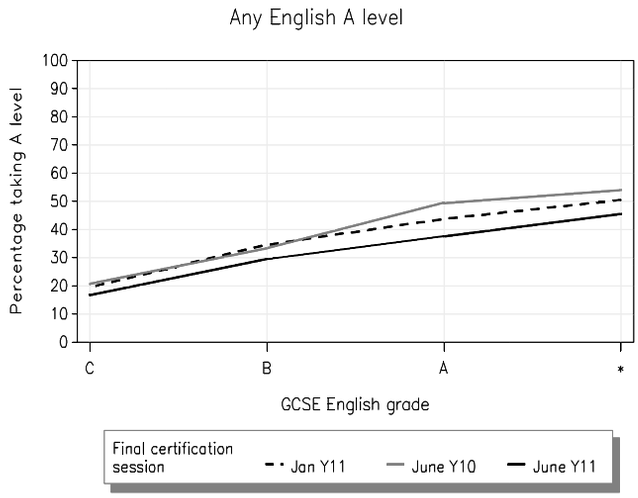
Figure 3: Maths 2009 – Percentage achieving key grades, by KS2 level and first certification session



		First certification session	*KS2 Maths level*			
			2	*3*	*4*	*5*
Grade A	First grade	Pre-June Y10	0.0	0.2	0.7	10.5
		June Y10	0.0	0.1	1.9	40.9
		Jan Y11	0.0	0.1	1.8	33.3
		June Y11	0.1	0.4	8.0	54.8
	Final grade	Pre-June Y10	0.0	0.5	5.8	36.9
		June Y10	0.2	0.5	4.7	51.3
		Jan Y11	0.0	0.3	2.9	38.6
		June Y11	0.1	0.4	8.0	54.8
Grade C	First grade	Pre-June Y10	0.0	2.1	25.5	74.6
		June Y10	0.6	4.8	37.9	89.8
		Jan Y11	1.2	12.5	52.7	92.4
		June Y11	2.8	16.5	68.7	97.8
	Final grade	Pre-June Y10	1.7	19.3	72.5	97.2
		June Y10	4.5	20.7	71.2	97.8
		Jan Y11	4.6	24.2	68.9	96.6
		June Y11	2.8	16.5	68.7	97.8

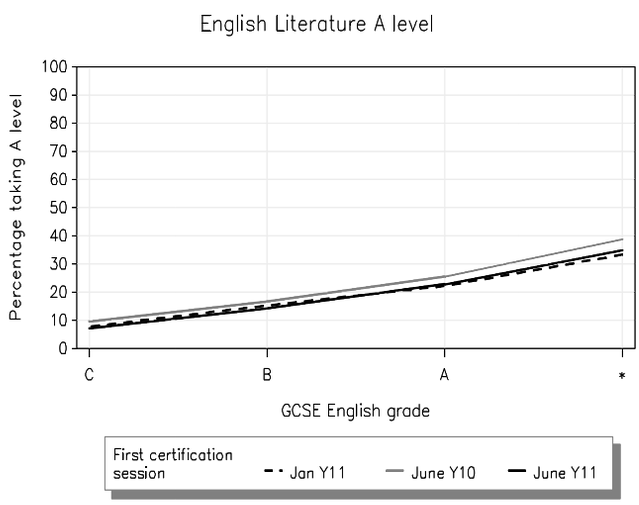
Figure 4: Maths 2011 – Percentage achieving key grades, by KS2 level and first certification session

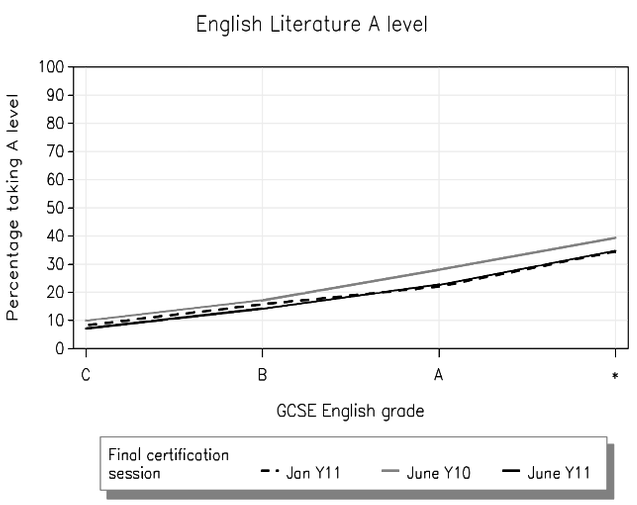




	GCSE session	*GCSE grade*			
		C	*B*	*A*	*A**
First	June Y10	19.8	34.3	48.2	53.2
	Jan Y11	19.0	34.3	44.0	49.9
	June Y11	16.7	29.4	37.4	45.5
Final	June Y10	20.7	33.2	49.4	54.0
	Jan Y11	19.4	34.5	43.7	50.4
	June Y11	16.7	29.4	37.4	45.5

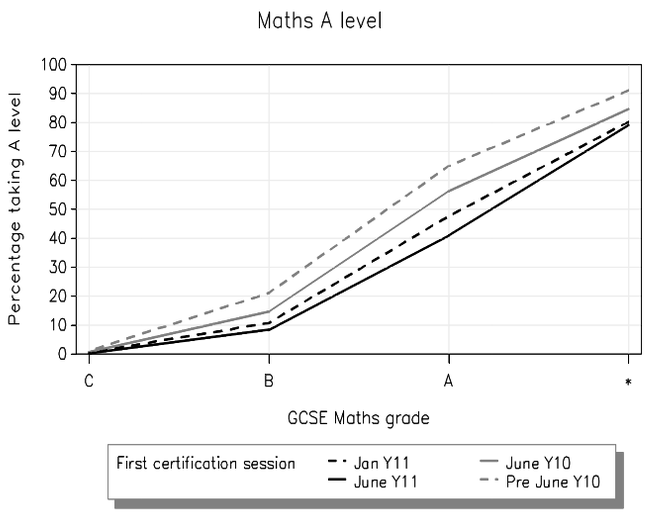
Figure 5: Percentage of A level candidates taking any English A level, by GCSE English grade and certification session (first and final)

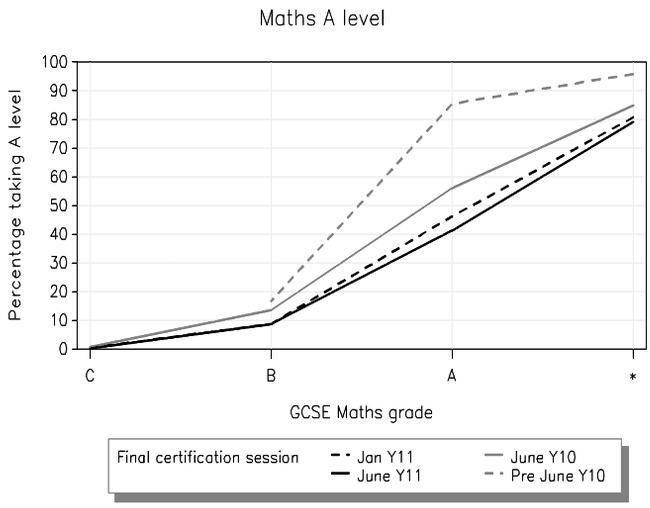




	GCSE session	*GCSE grade*			
		C	*B*	*A*	*A**
First	June Y10	9.7	16.6	25.5	38.7
	Jan Y11	7.7	15.1	22.2	33.3
	June Y11	7.1	14.1	22.8	34.8
Final	June Y10	10.0	17.2	27.9	39.2
	Jan Y11	8.2	15.7	22.0	34.5
	June Y11	7.2	14.2	22.7	34.8

Figure 6: Percentage of A level candidates taking English Literature A level, by GCSE English grade and certification session (first and final)





	GCSE session	*GCSE grade*			
		C	*B*	*A*	*A**
First	Pre-June Y10	1.1	21.1	64.9	91.0
	June Y10	0.8	14.6	56.2	84.5
	Jan Y11	0.5	10.7	47.6	80.2
	June Y11	0.4	8.4	40.8	79.0
Final	Pre-June Y10	0.0	16.7	85.4	95.8
	June Y10	0.7	13.6	56.0	84.9
	Jan Y11	0.3	8.7	46.2	80.7
	June Y11	0.4	8.7	41.2	79.0

Figure 7: Percentage of A level candidates taking Maths A level, by GCSE Maths grade and certification session (first and final)

Figure 7 presents the percentage of candidates achieving each of grades A* to C at GCSE Maths going on to take Maths A level. For both first certification session and final certification session, candidates who took the GCSE in June of year 11 were the least likely to go on to take Maths A level (for all GCSE grades). Those certificating prior to June of year 10 were most likely to take an A level in the subject.

These figures suggest that there is no detrimental effect on A level uptake of certificating early in the GCSE in the subject. In fact, there is some evidence that candidates who certificate early are more likely to go on to study at A level. This is the case for both the first and the final certification session.

A level performance

The indicators of A level performance used were the achievement of a grade C or better and achievement of a grade A or better in the relevant A level. The A level subjects investigated were English (any), English Literature, Maths, Biology, Chemistry and Physics. For the Science subjects candidates were categorised by when they certificated in GCSE Maths, to investigate the hypothesis that those who certificated early in Maths may be disadvantaged in Science A levels because of their mathematical content.

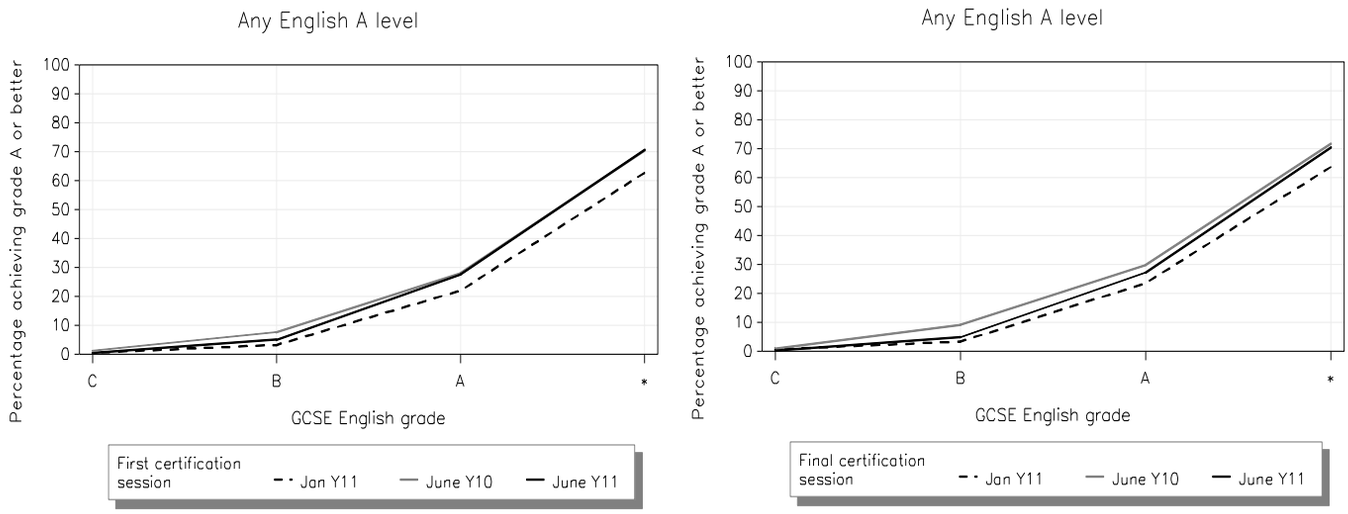
As with A level uptake, separate analyses were undertaken for first certification session and final certification session. Candidates certificating early may be disadvantaged because their GCSE study was rushed and did not prepare them fully for further study (even if they re-sit). The gap in studying for those whose final certification was early may also disadvantage some candidates because they have forgotten what they learnt. Candidates were also categorised by the grade they achieved in the subject at GCSE, as higher achieving GCSE candidates are more likely to perform better at A level. It should also be noted that for the Science A levels, GCSE grade refers to the grade achieved on the most advanced Science taken by the candidate. Thus, for those not taking separate Science GCSEs, if the candidate took the Additional Science qualification then the grade refers to this subject. If not, then the grade refers to that achieved in Core Science.

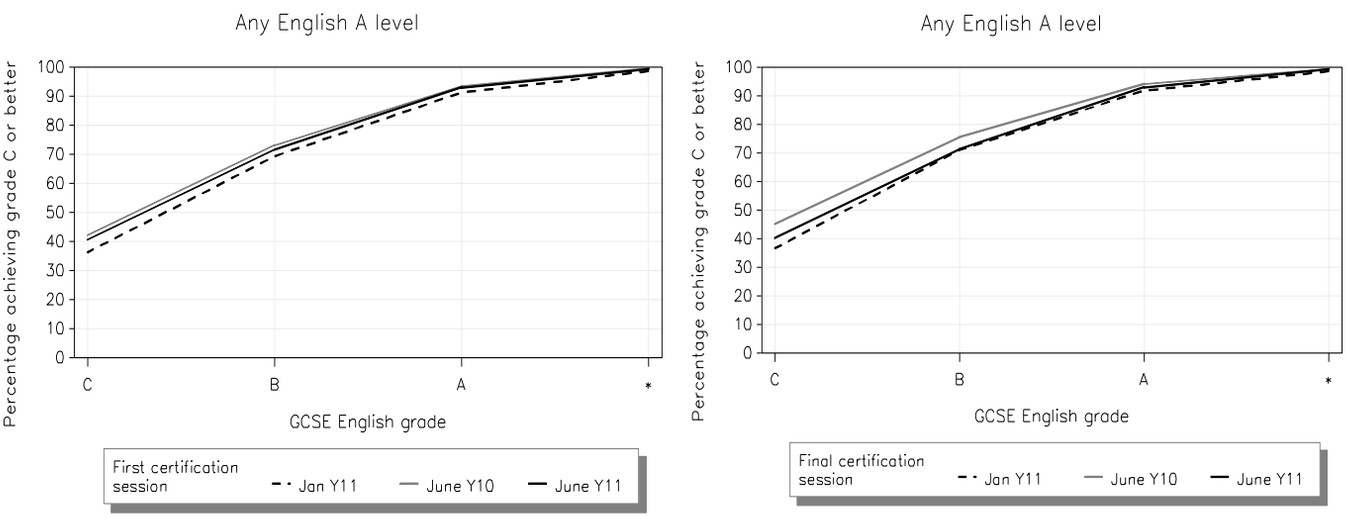
Table 9 presents the number of matched candidates taking the A level and the percentages achieving grade A or better and grade C or better.

Figures 8 to 13 present the percentages of candidates achieving at least a grade A and at least a grade C in an A level subject, by their GCSE grade in the subject and the GCSE certification session in the same or a

Table 9: Number of matched A level candidates and percentages achieving grade A or better or grade C or better

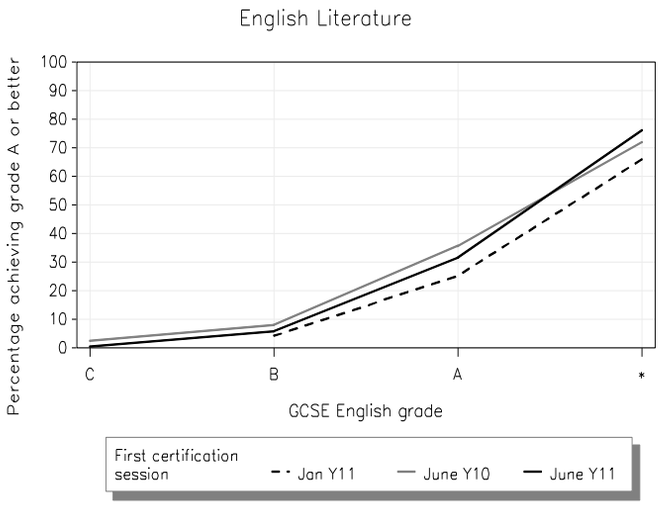
GCSE Subject	*A level Subject*	*Matched candidates*	*% getting grade A or better*	*% getting grade C or better*
English	Any English	70,258	23.0	79.8
English	English Literature	39,236	30.2	81.7
Maths	Maths	54,146	43.2	81.5
Maths	Biology	42,831	27.9	73.1
Maths	Chemistry	32,545	33.6	78.4
Maths	Physics	21,890	31.8	72.7

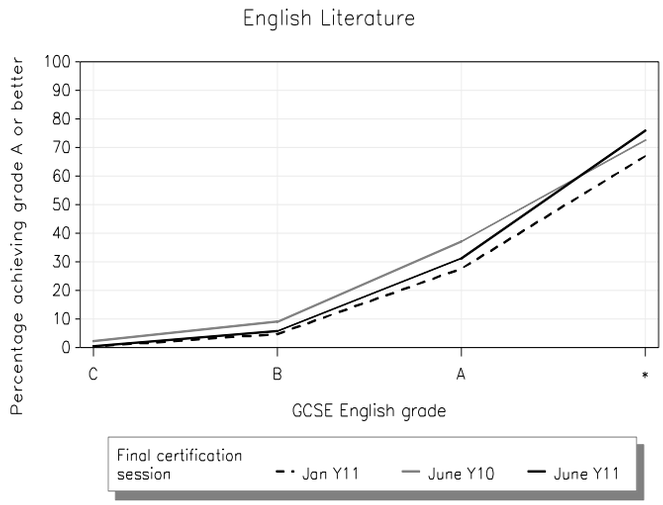


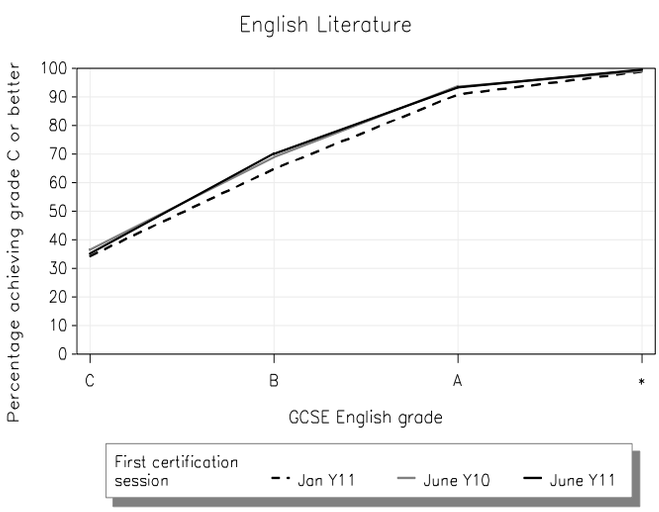


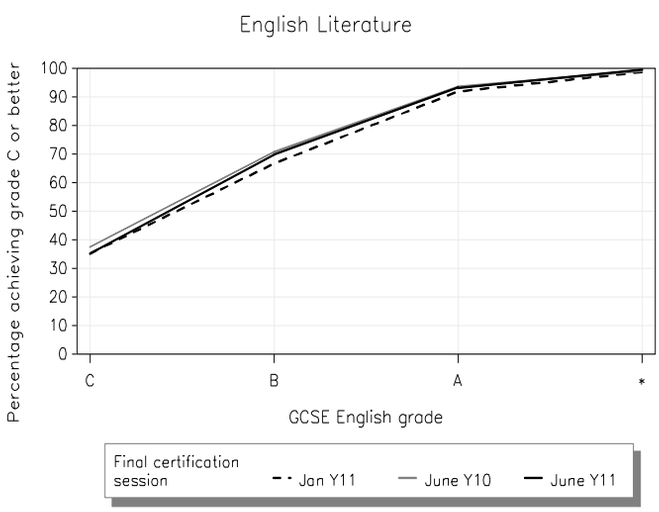
		GCSE session	*GCSE grade*			
			C	*B*	*A*	*A**
Grade A	First	June Y10	1.2	7.6	28.0	70.4
		Jan Y11	0.5	3.2	21.9	62.6
		June Y11	0.5	5.0	27.5	70.6
	Final	June Y10	1.1	9.1	29.7	71.7
		Jan Y11	0.6	3.3	23.5	63.6
		June Y11	0.5	4.9	27.2	70.4
Grade C	First	June Y10	42.1	73.1	93.4	99.4
		Jan Y11	36.2	69.2	91.2	98.5
		June Y11	40.6	71.6	93.1	99.3
	Final	June Y10	45.2	75.4	94.1	99.4
		Jan Y11	36.5	71.0	91.7	98.5
		June Y11	40.2	71.3	93.0	99.3

Figure 8: Percentage of candidates taking any English A level achieving key grades, by GCSE English grade and certification session (first and final)



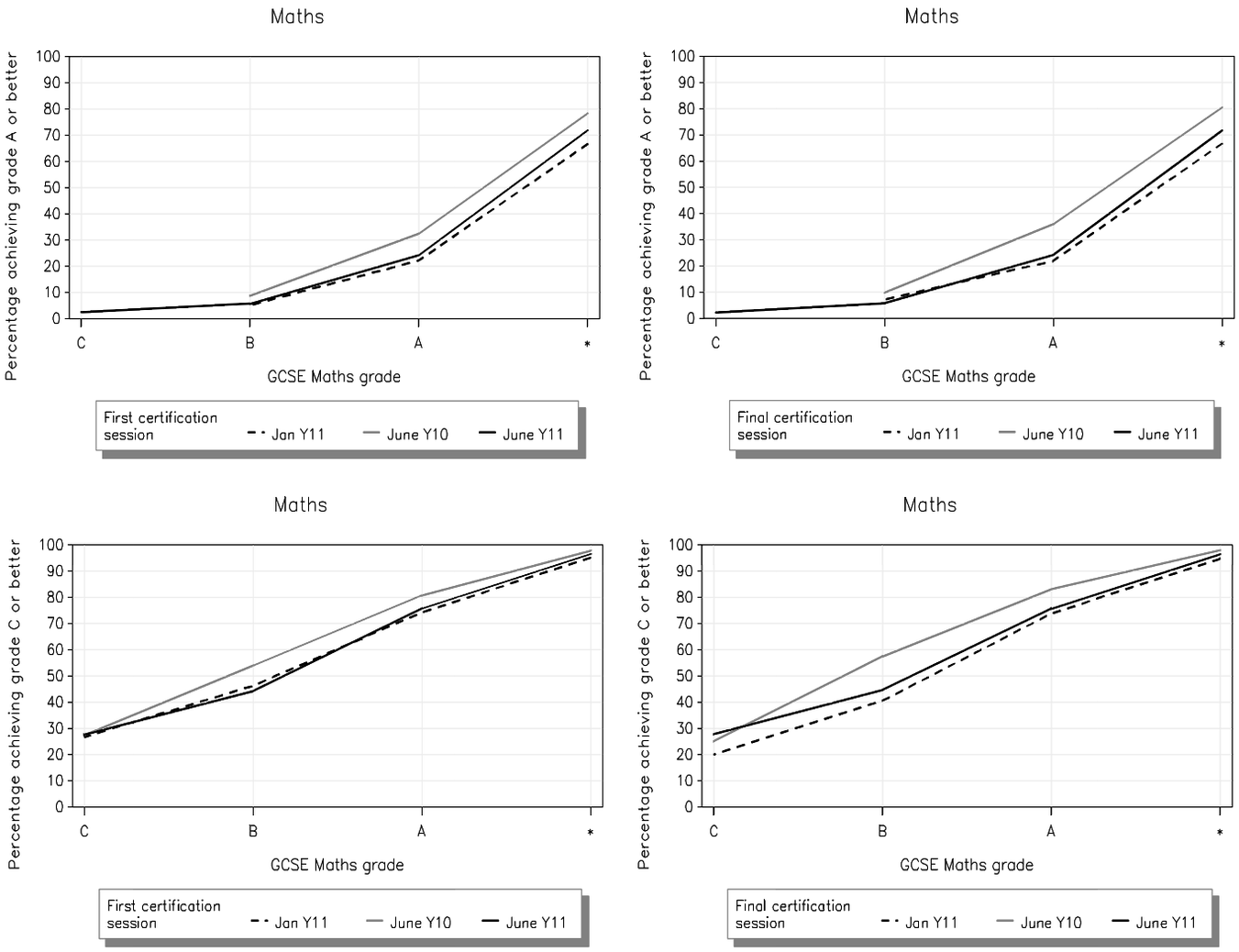






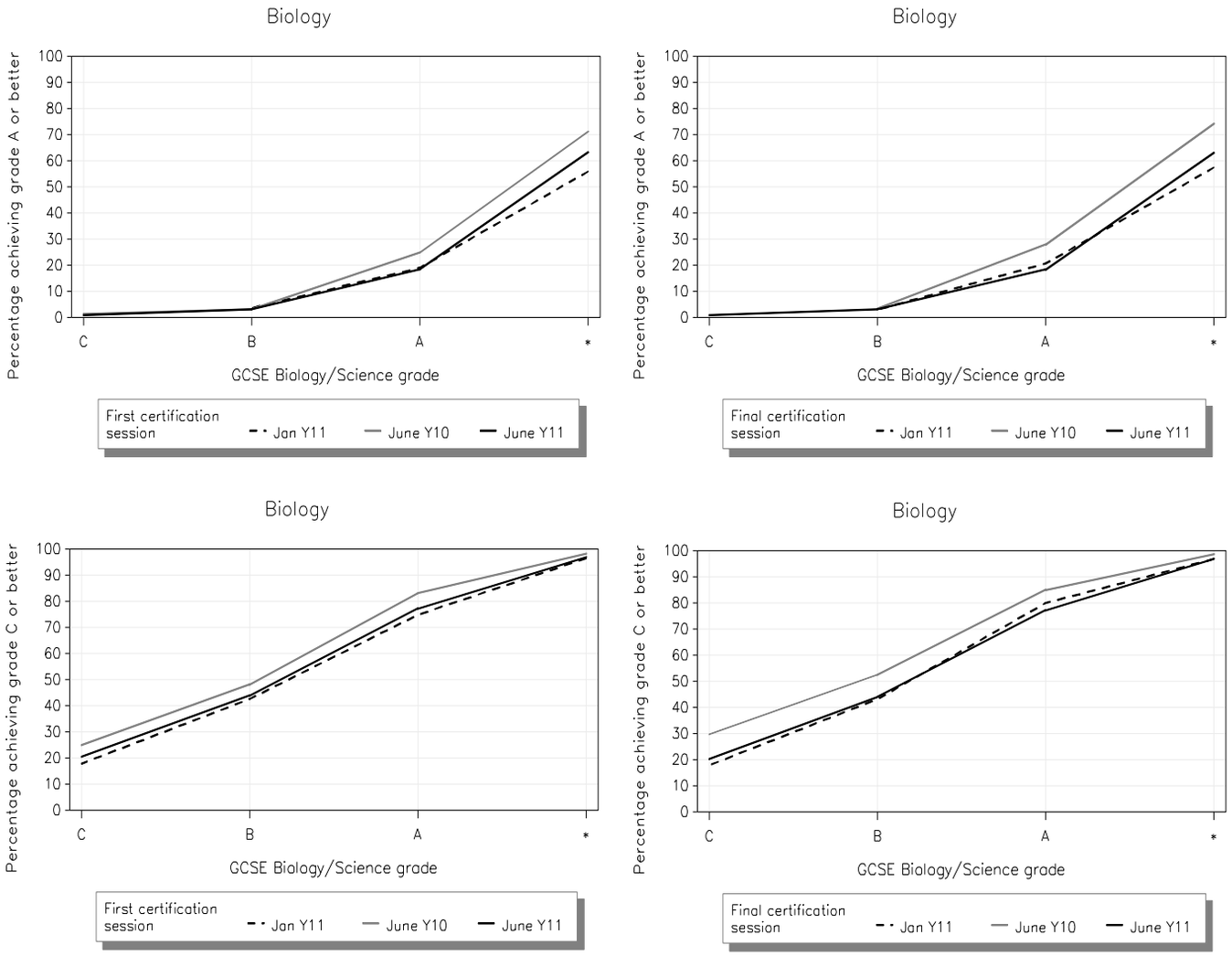
		GCSE session	*GCSE grade*			
			C	*B*	*A*	*A**
Grade A	First	June Y10	2.5	8.0	35.7	71.9
		Jan Y11	0.0	4.1	25.2	65.8
		June Y11	0.6	5.8	31.5	76.1
	Final	June Y10	2.3	9.1	37.1	72.5
		Jan Y11	0.3	4.7	27.5	66.9
		June Y11	0.5	5.7	31.2	75.9
Grade C	First	June Y10	36.6	68.9	93.7	99.1
		Jan Y11	34.2	64.7	90.8	98.8
		June Y11	35.3	70.2	93.3	99.4
	Final	June Y10	37.5	70.7	93.5	99.0
		Jan Y11	35.1	66.6	91.8	98.6
		June Y11	35.1	69.8	93.2	99.5

Figure 9: Percentage of candidates taking English Literature A level achieving key grades, by GCSE English grade and certification session (first and final)



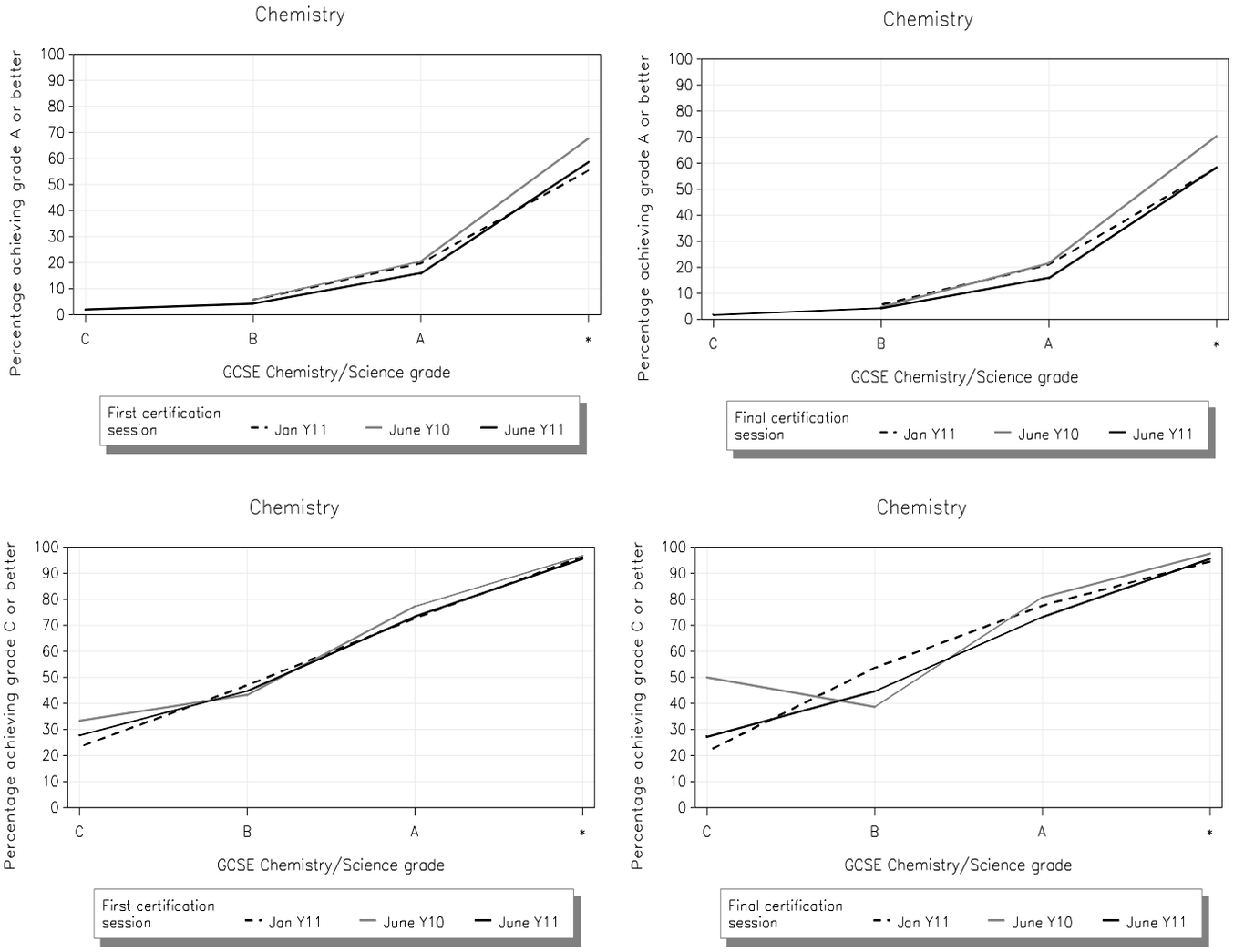
		GCSE session	*GCSE grade*			
			C	*B*	*A*	*A**
Grade A	First	June Y10	0.0	8.9	32.5	78.2
		Jan Y11	0.0	5.1	22.1	66.6
		June Y11	2.6	5.7	24.2	71.8
	Final	June Y10	0.0	9.7	35.9	80.4
		Jan Y11	0.0	7.1	21.9	66.7
		June Y11	2.2	5.8	24.2	71.6
Grade C	First	June Y10	27.3	53.9	80.7	97.7
		Jan Y11	26.7	46.2	74.2	95.0
		June Y11	27.7	44.2	75.7	96.5
	Final	June Y10	25.0	57.4	83.1	97.9
		Jan Y11	20.0	40.5	73.7	94.6
		June Y11	27.7	44.6	75.6	96.4

Figure 10: Percentage of candidates taking Maths A level achieving key grades, by GCSE Maths grade and certification session (first and final)



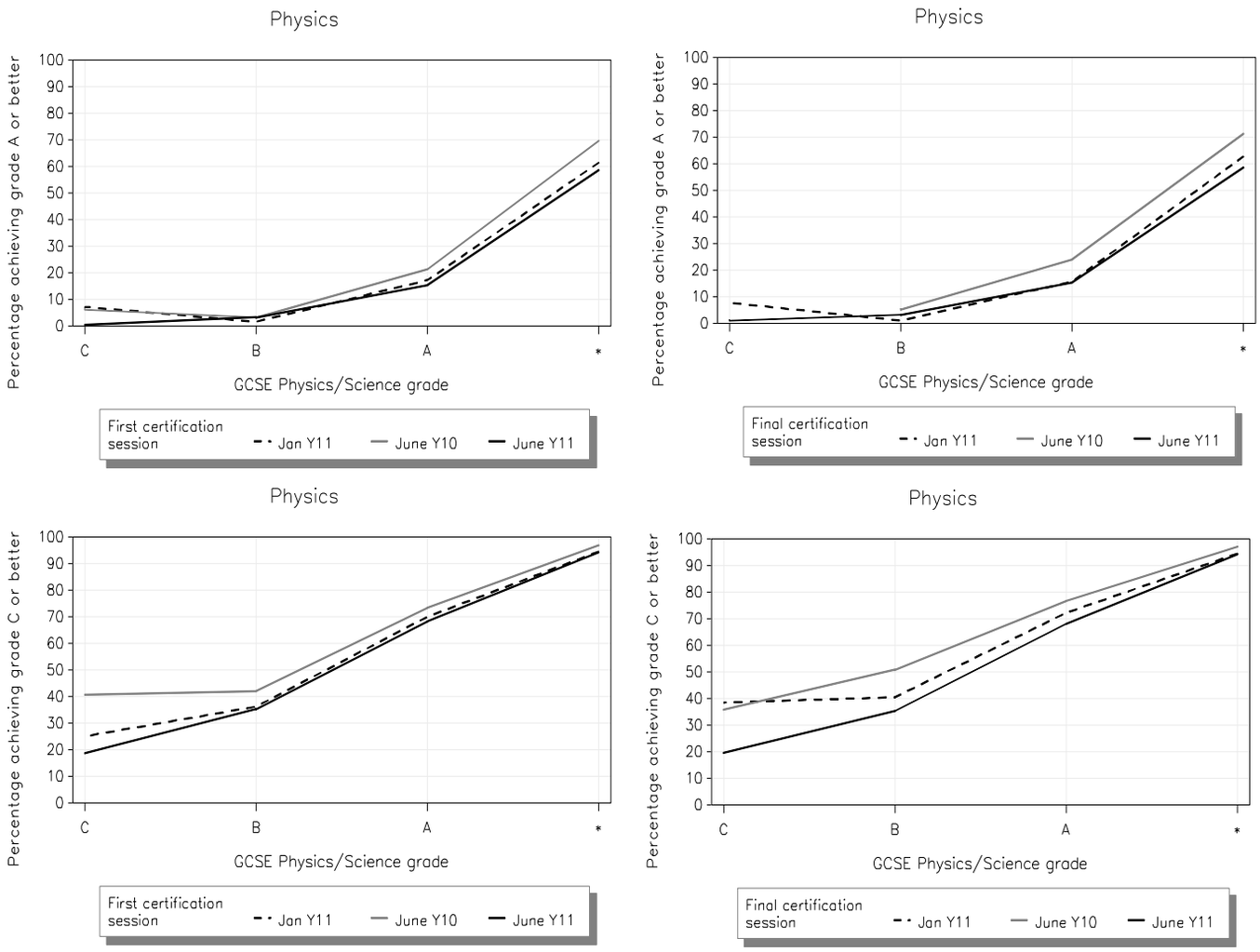
		GCSE session	*GCSE grade*			
			C	*B*	*A*	*A**
Grade A	First	June Y10	1.5	3.2	24.9	71.1
		Jan Y11	0.0	3.5	19.0	55.8
		June Y11	1.0	3.1	18.4	63.2
	Final	June Y10	0.0	3.4	27.9	74.2
		Jan Y11	0.0	2.9	20.5	57.4
		June Y11	1.0	3.1	18.4	63.0
Grade C	First	June Y10	25.0	48.2	83.1	98.1
		Jan Y11	17.8	42.5	74.7	96.3
		June Y11	20.5	44.0	77.3	96.9
	Final	June Y10	29.6	52.5	85.0	98.5
		Jan Y11	17.8	43.1	79.9	96.7
		June Y11	20.4	44.0	77.2	96.9

Figure 11: Percentage of candidates taking Biology A level achieving key grades, by GCSE Biology/Science grade and GCSE Maths certification session (first and final)



		GCSE session	*GCSE grade*			
			C	*B*	*A*	*A**
Grade A	First	June Y10	0.0	5.7	20.5	67.7
		Jan Y11	0.0	5.7	19.7	55.4
		June Y11	2.1	4.2	15.9	58.6
	Final	June Y10	0.0	4.7	21.5	70.3
		Jan Y11	0.0	5.6	21.0	58.2
		June Y11	1.9	4.3	16.0	58.4
Grade C	First	June Y10	33.3	43.3	77.3	96.7
		Jan Y11	23.3	46.9	72.6	96.3
		June Y11	27.8	44.7	73.3	95.5
	Final	June Y10	50.0	38.6	80.7	97.5
		Jan Y11	21.4	53.6	77.5	94.4
		June Y11	27.3	44.7	73.2	95.5

Figure 12: Percentage of candidates taking Chemistry A level achieving key grades, by GCSE Chemistry/Science grade and GCSE Maths certification session (first and final)



		GCSE session	*GCSE grade*			
			C	*B*	*A*	*A**
Grade A	First	June Y10	6.3	3.2	21.3	69.6
		Jan Y11	7.1	1.5	17.2	61.3
		June Y11	0.5	3.3	15.3	58.6
	Final	June Y10	0.0	5.3	24.0	71.2
		Jan Y11	7.7	1.0	15.7	62.7
		June Y11	1.1	3.2	15.3	58.5
Grade C	First	June Y10	40.6	41.9	73.2	96.7
		Jan Y11	25.0	36.2	69.9	94.4
		June Y11	18.8	35.3	68.2	94.2
	Final	June Y10	35.7	50.8	76.7	97.0
		Jan Y11	38.8	40.4	72.2	94.4
		June Y11	19.5	35.2	68.0	94.2

Figure 13: Percentage of candidates taking Physics A level achieving key grades, by GCSE Physics/Science grade and GCSE Maths certification session (first and final)

related subject. The left-hand graphs are for the *first* certification session for the GCSE and the right-hand graph for the *final* certification session

The figures show results that are very similar for all subjects. Candidates certificating in the GCSE in June of year 10 were almost always the most likely to get a grade A or better or a grade C or better in the A level (for each level of prior attainment). This was true for both the first and the final certification sessions at GCSE. Otherwise, for most subjects there was very little difference in the percentage achieving the key grades or better for those certificating in January of year 11 and those certificating in June of year 11. The only exception was in relation to any English A level, with those certificating in January of year 11 less likely to get a grade A or better than those certificating in June of year 11.

This suggests there is very little evidence that candidates who certificate early at GCSE are disadvantaged in their performance at A level, either through poor preparation because of cramming the work into a shorter period of time, or through having a gap between GCSE and A level study. It should be noted that we can only infer that candidates who certificate early and do not re-sit are, in fact, taking a break from studying the subject. It may be that many of them continue their study, either through non-GCSE qualifications such as the Free Standing Maths Qualification or through non-certificating courses.

Discussion

Analysis of the entry patterns in GCSEs has shown an upsurge in early certification since the scrapping of the KS3 tests. This is particularly the case in English and Maths where a high percentage of candidates in 2011 certificated in January of year 11. The motivation for entering so many candidates early is not known, but it may be an attempt to bank a good grade in a subject so that other subjects can be focused on in the summer of year 11.

An analysis of entries by school type showed that candidates in comprehensive, academy and secondary modern schools were more likely than those in independent schools to certificate early in either English or Maths. It is not known why this is the case but it highlights the influence of school strategy on early entry patterns, with some schools entering whole classes early, perhaps in order to try and get the all-important grade Cs for league table purposes and other accountability measures. Other schools are likely to be more sensitive to the aptitudes of individual candidates, whilst others still may not have any candidates entering early.

The increase in early certification in recent years begs the question: are candidates who certificate early performing below their potential? In the data on GCSE performance in English and Maths there was some evidence of this. Candidates who certificated early were (mostly) less likely to achieve a grade A or better than those who certificated at the expected time, both in terms of first grade and final grade. Candidates who certificated early were also less likely to achieve at least a grade C in their first certification. However, when final grade was taken into account, there was very little difference in the likelihood of achieving a grade C or better between the different certification sessions. Thus, it seems that high performing candidates who certificate early are more likely to perform below their potential, but for candidates of lesser ability there is apparently no particular disadvantage of early certification. This outcome may be because of the focus on attaining grade C at GCSE; higher ability candidates who certificate early would not find it difficult to achieve a grade C so are given less attention by teachers focussing on getting less able candidates up to grade C. However, the higher ability candidates may not yet be ready to get up to grade A and so are performing below their potential.

In terms of A level uptake, candidates certificating at the expected time generally had the lowest probability of going on to take the subject at A level, after accounting for GCSE grade in the subject. So there is no evidence that having a break from studying for a subject leads to loss of interest or lack of confidence in continuing study.

Finally, there was little evidence that the performance at A level of candidates who certificated early in the GCSE subject was worse than that of candidates who certificated at the expected time, after accounting for prior attainment. Across all subjects (and GCSE grade), candidates who certificated for GCSE in June of year 10 had a higher probability of achieving a grade A or better or a grade C or better at A level than those certificating at the expected time. Only in relation to any English A level were some candidates seemingly disadvantaged by certificating early (in terms of the probability of achieving a grade A or better).

It should be noted that there is an issue with using the GCSE grade as prior attainment in the A level analysis. Figures 1 and 3 showed that candidates in the 2009 cohort certificating early at GCSE were more likely to perform below their potential. Thus, for these candidates, the GCSE grade used in the A level analyses may not reflect their 'true' ability in the subject. Now, assume that the likelihood of going on to take an A level is dependent (to some extent) on a candidate's true ability in the subject. Then, candidates who certificated early and received, for example, a grade B, but whose 'true' ability was a grade A would have a higher probability of going on take an A level than equivalent candidates who certificated at the usual time and received a grade B (a true reflection of their ability). This would then have the effect of artificially inflating the percentage of those certificating early and getting a grade B going on to take an A level (or achieving a particular A level grade or better). This might explain why candidates certificating at the expected time were apparently the least likely to go on to take an A level and (to some degree) less likely to achieve a particular grade or higher at A level.

We should also be careful about drawing too many firm conclusions about the effect of early certification on performance. This analysis has shown an association between certification session and GCSE grade, but this does not necessarily imply causation. It may be that candidates who certificate early are less likely to do well in the subject for a reason unrelated to when they take the exam.

When following up students' performance at A level it was only possible to look at the cohort of students who finished KS4 in 2009. This was before the real upsurge of early entry at GCSE. It may therefore be that most of the students taking the GCSE a year early had a particular aptitude for the subject and were therefore likely to perform well at A level. It would be interesting to see how well the GCSE 2011 cohort does in comparison, using data from the 2013 NPD when available.

Finally, it is also worth considering what the impact of the changes to GCSEs will have on the amount of early certification. Candidates starting GCSEs in September 2012 will have to take all exams at the end of the course, instead of being able to take some exams earlier in the

course. However, this does not prevent early entry, as it will still be possible to take the final exams in the summer of year 10 (from June 2014 onwards). This may mean fewer candidates certificate early because there are fewer opportunities to do so. However, it is also possible that more candidates will certificate early, because it will no longer be possible to take individual units early and then re-sit them later in the course as required. Instead, candidates may sit all their exams early, so that they then have the opportunity to re-sit if required.

References

Department for Education (2011). *Early entry to GCSE examinations*. Available online at: https://www.education.gov.uk/publications/eOrderingDownload/DFE-RR208.pdf (Accessed 14 May 2012).

Gill, T. (2010). *How old are GCSE candidates? Statistics Report Series No.20*. Cambridge: Cambridge Assessment.

Vidal Rodeiro, C. L. & Nadas, R. (2012). Effects of modularity, certification session and re-sits on examination performance. *Assessment in Education: Principles, Policy and Practice*. **19**, 4, 411–430.

Reaching for the A*: Exploring the extent and effect of resitting at A level

Tom Sutch and **Frances Wilson** Research Division

Introduction

The introduction of Curriculum 2000 changed the traditional linear structure of A levels to a modular structure, and introduced an integrated AS level qualification, comprising half of the modules and set at the standard expected of A level students after one year of study (Tomlinson, 2002). This reform afforded candidates opportunities to resit individual modules to improve their grades, using the best results obtained in each module to count towards the A level. However, there has been frequent criticism that this has led to a resit culture, with students resitting modules multiple times until they achieve their desired grade, leading to fears that students may be achieving high grades at A level by resitting. In November 2006 changes to A level specifications were agreed. These changes included the introduction of the new A* grade. The new specifications were first taught from September 2008, and the first of the new A* grades were awarded in June 2010. The new A* grade was introduced to differentiate between the highest achieving candidates so that universities could select the best candidates. The A* is awarded to candidates who fulfil two criteria. First, candidates must achieve at least 80% of Uniform Mark Scale (UMS) marks overall (i.e. an A grade), and must achieve 90% or more of UMS marks at A2. In an open letter to schools, Ofqual (2010) stated that this structure aimed to reward consistently good performance throughout the A level, and to reward exceptional performance at A2. This structure also aims to reduce the incentive to resit AS modules, because a lower threshold is needed at AS level than at A2.

This study investigated resitting patterns in five OCR GCE A level specifications with a particular focus on high achieving students who achieved the new A* grade.

Background

Despite the fact that the introduction of the opportunity to resit units, as part of modularisation, represented a major reform in the structure of A levels, there has been relatively little literature from government or the qualifications regulator on the intended purposes of allowing resitting at a unit level.

One obvious rationale, given the high-stakes nature of the A level, is that resitting gives students a chance to set the record straight if they performed less well than expected, wished for or 'deserved' following a bad day with a particular examination[1]. In 2007, Ken Boston, then Chief Executive of QCA, stated that 'candidates deserve a chance to demonstrate their ability if they failed to do so the first time' (MacLeod, 2007). This argument would apply to linear A levels too, albeit at a less fine-grained level. In a discussion of modularity in A levels, Dearing (1996) noted that one of the reported advantages of modular syllabuses is that they give 'an opportunity to resit a module and achieve, on merit, a better result through additional work' (p.90). This is a broader argument that allows for the role of unit assessment to be formative as well as summative, and arguably (depending on what is meant by 'additional work') accounts for the reinforcement of learning through coverage of related material in later units; however, taking advantage of maturational effects in this way could be seen by some as dubious. Gray (2002) argued that resitting was a legitimate and integral part of a modular assessment regime, and ultimately improved student attainment through mechanisms of feedback, multiple opportunity and motivation. When QCA removed a short-lived restriction on the number of resits per A level unit, it was for practical rather than ideological reasons, with the justification that results would not be greatly affected due to the low numbers likely to take advantage of multiple resitting (QCA, 2003b).

Poon Scott (2011) argued that the resit policy had unintended consequences, namely students and teachers devising elaborate resit strategies and an increased focus on the assessment process itself, which had a negative effect on student learning. There is a view commonly held, including by teachers (de Waal, 2009; Higton *et al*., 2012), that a 'resit culture' has developed: because there is no penalty attached within the overall A level mark[2] for resitting units an unlimited number of times; students are entered for unit exams early to give them more test experience; and they resit as often as necessary to gain their desired

1. Although with a broader purpose than the special consideration provisions for adjusting marks for candidates who, for example, are ill on the day of the examination or recently bereaved.
2. However, there are other penalties, as resitting has implications for cost, loss of teaching time, and some universities look less favourably on marks gained through resitting. In practice there are only a few opportunities to resit.

grade, effectively by attrition. Warnings and criticisms of a 'resit culture' date back to the introduction of the current A levels under the Curriculum 2000 reforms (QCA, 2003a). In this view, AS units are favoured for resitting as they are easier than A2 units but contribute equally to the overall A level. However, as Gill and Suto (2012) describe, there are more opportunities available to resit AS units so higher resit rates are not unexpected. Additionally, Gill and Suto (2012) found that the situation was nuanced: while students and teachers were strategic in deciding whether and what to resit, very few students said they treated the first sitting of unit exams as a practice.

The inherent unreliability and measurement error of the assessment process means that many candidates will increase their scores, and sometimes grades, on a resit. However, when regression to the mean is taken into account (Smith and Smith, 2005) we should expect a clear increase in scores: essentially, resitting candidates are more likely to be those that have achieved below their true score at the first sitting, and randomness alone will tend to make their average increase. Wheadon (2010) simulated the impact of resitting a typical AS unit and found that, as expected, the rate of false negatives (candidates achieving less than their true ability) fell with each successive attempt at a test, but the rate of false positives (candidates outperforming their true grade) rose[3]. In addition, given that it is not possible to resit a unit until six months later, other effects such as maturation and coverage of related topics as part of other units may well contribute to an improvement.

In early investigations of resitting under the new Curriculum 2000 A levels, QCA (2003c) found that a substantial minority of students resat AS units once, and there was generally little resitting of A2 units. However, differences in A2 resitting were observed between subjects: 25–33% of candidates in Physics resat A2 units but candidates in French hardly resat. QCA attributed this difference to the common practice in science subjects of sitting the first A2 unit in the January session. Most students who did resit improved their marks. Further investigations were undertaken as part of a study of A level Mathematics (Matthews and Pepper, 2007) using awarding body data from six subjects. The study found that resit rates were higher in Mathematics than in the other subjects. Mathematics students were more likely to resit units multiple times than in other subjects, although there were also higher than average multiple resits in French, which the authors attributed to the benefits of maturation. There was very little multiple resitting of A2 units, except in Mathematics. The effect of resitting on the proportion of candidates awarded grade A was also estimated by calculating a notional A level grade using the AS result at the end of the first year of study[4]. This showed an increase of between 2.3 and 7.8 percentage points due to resitting.

Gill (2010) found that resitting had more of an effect on lower grades, essentially because there were more opportunities to increase a grade (from a grade C at first sitting, one can improve to a B, A or A*; from a grade A at first sitting, only an A* would be an improvement). Similarly, Wheadon (2010) found that resitting had less of an impact on A grade candidates than lower grades, although this analysis dates from before the introduction of the A*.

Empirical results from AQA A levels in 2011 (CERP, 2012) showed that the candidates achieving the best grades tended to resit the fewest units, and that resitting brought most benefit to the candidates with middle grades (B, C and D). We might expect, therefore, that candidates who have been awarded an A* in their A level would have resat fewer units than average, based on their high ability. QCA (2007) suggested that the design rules for A* should encourage a reduction in resitting, but a student or centre would need to feel confident that the A grade threshold was achievable in order to risk the lower level of AS resitting.

However, there has been feedback from teachers (QCA, 2007) that there are exceptional candidates who resit units in order to maximise their overall UMS mark (even in pursuit of a few extra marks) because of requirements for entry into higher education. This is not directly related to the A* grade but is an artefact of the availability of detailed mark scores. Williams (2009, pp.152–154) reported that many students aim for marks of 90% UMS in AS, in the belief that this is an unwritten requirement for Oxbridge and the top Russell Group universities. Gill (2010) observed a similar effect in GCSE units which some candidates resat, after achieving an A* unit grade, in order to obtain a 'good' A* and boost their overall grade.

Aims

This study aimed to compare the resitting patterns of students achieving the new A* grade with less highly achieving students across five contrasting A level subjects. In particular, we investigated two main areas: the extent of resitting across different grades, and the effect of resitting on the final grade and marks. The extent of resitting was examined to determine whether high achieving A* students were more or less likely to resit units than other candidates. Patterns of resitting were investigated with respect to the number of units resat by candidates achieving each grade and the probability of a candidate resitting given a previous mark. Since candidates may base their decision to resit a unit on the grade that they expect to achieve, we also examined the extent of resitting based on a student's forecast grade. It is also important to consider the effect of resitting, to determine whether students seem to be remediating an anomalous performance or capitalising on chance to improve their grade. We investigated the change in mark at the unit level, and the change in overall grade achieved by resitting. Since the A* is awarded using a different rule from other grades (which was designed partly to reduce the effect of resitting), we then focussed on whether resitting resulted in different patterns of achievement for candidates awarded an A* or A grade.

Methods

Five contrasting OCR A level specifications were analysed in this study:

- Chemistry A: H434
- History: H506
- English Literature: H471
- French: H475
- Art and Design (Fine Art): H561

All specifications comprised four units, except Chemistry which had six units.

3. Wheadon's simulation results do not appear to be wholly consistent with the method described, which states that candidates would only resit if they achieved *below* their true grade. Hence, after repeated simulated tests one would expect the percentage of candidates with a grade *equal* to their true grade to increase or stay the same. However, the reported results show this percentage actually decreasing at the expense of false positives. Wheadon's false positive rate may therefore be overstated.

4. This could include some resitting of AS units first taken in January.

Three datasets were assembled for each of the specifications, containing all the unit-level results of candidates who certificated an A level for these specifications in 2012, along with the final A level grade. Note that candidates who were awarded an AS qualification only were excluded. The first dataset contained candidates' highest result for each unit (which would be the result cashed in), the second contained results from each candidate's first sitting of each unit, and the third contained marks from all sittings of each unit. Checks were performed to ensure that each candidate had the correct number of units.

When comparing performance at first sitting of a unit with that of the best performance ('before' and 'after' resit) we have implicitly assumed that the results obtained at the first sitting would reflect the situation if resits were not permitted. However, this may not be the case, especially if (as some have claimed) many students treat the first sitting as a practice.

Extent of resitting

Table 1 shows the percentage of candidates who resat one or more units for each subject. There is a wide variation in resit rates between subjects, with over half of candidates resitting at least one unit in Chemistry, History and French, but only 8% of candidates in Art and Design. The particularly high rate for Chemistry is likely to be because it is the only six-unit A level specification considered.

Equally clear is the difference in resit rates between grades. In each subject, candidates awarded an A* were less likely to have resat than those awarded an A, and much less likely than the average.

Table 1: Rates of resitting by subject

Subject	*Total candidates*	*% candidates resitting*		
		*A**	*A*	*All*
Chemistry	18149	21.9	55.8	76.4
English Literature	10120	14.9	27.8	37.7
History	10810	36.4	44.2	59.5
French	1095	9.2	40.6	55.0
Art & Design	3719	1.7	4.7	7.7

Average units resat by grade

Figure 1 shows the average number of units resat (one or more times) by candidates awarded different A level grades. The size of each point represents the number of candidates awarded each grade.

There is a clear trend for the rate of resitting to be highest in the middle of the grade distribution, with a slight drop for the lowest grades, but with candidates awarded an A* hardly resitting at all. The highest resit rate for A* candidates was in History (0.46 units).

In general, candidates resat more AS units than A2, with the exception of Art and Design which had very low resit rates across both halves of the A level. The low rate of resitting in Art and Design is likely to be due to the nature of the assessment (most of the marks are gained via a coursework portfolio), and the limited opportunities to resit, because no assessment is available in the January sessions.

Resitting A2 units was more common in Chemistry than other subjects. Due to the six-unit specification, many candidates took one A2 unit in the January session of their final year, thus giving an opportunity to resit in June should this be desired. By contrast, for

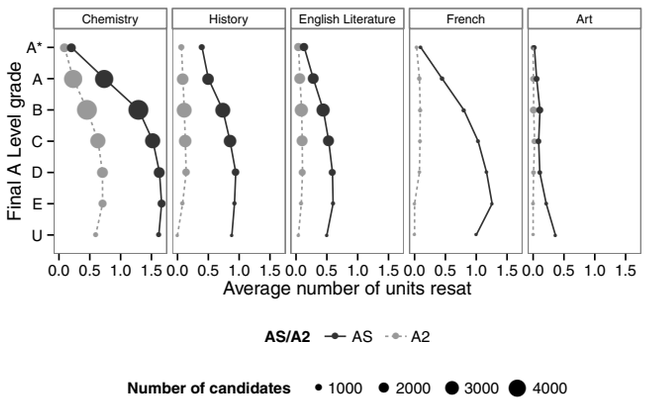


Figure 1: Average units resat by grade

History, English Literature and French, most candidates took all their A2 units in the June sessions.

Resitting decisions given a particular unit mark

One possible explanation for the low resitting rates for the candidates of highest ability, as observed in Figure 1, may simply be that because they would tend to get the best marks at a unit level, they would have less need to resit them. To examine this, plots are presented in Figure 2 showing resitting behaviour for all candidates *given the original mark obtained in each unit*, and similarly for A* and A candidates only.

Each point shows, for a given percentage UMS, the proportion of candidates obtaining that score (on their first sitting of the unit) who decided to resit the unit. It should be borne in mind that each point represents a different number of candidates: to deal with this to some extent, lines have been generated using Loess smoothing[5] weighted by the number of candidates. This gives an indication of the scores that different groups of candidates were comfortable with.

The plot has been cropped to focus on the area of interest (60–100% UMS). It is technically possible for candidates to obtain an A* by scoring an average of 60% UMS on their AS units and 100% on their A2 units, but this would be a risky strategy, so if candidates were aiming for an A* one would expect them to resit given marks of, say, below 60–80%. On the other hand, if candidates resat an AS unit after scoring marks of over 80%, it might indicate that they were trying to gain insurance against harder A2 units, or resitting opportunistically (perhaps during their A2 year) because they found this particular unit relatively easy.

Only AS units are shown; resitting patterns were different in A2, as candidates have to score at least 90% UMS on average to be awarded an A*, leaving little room for error: with two equally weighted units, the lowest mark that could be scored on one of them is 80% (with 100% in the other). Whether or not a candidate resits given a particular mark is therefore closely linked to whether they go on to achieve an A*.

In general, the proportion of candidates resitting decreases as the original mark increases, as one would expect. There was little resitting in response to original marks of 80% or more for most subjects. However, History F962 and French F702 were exceptions: in the French unit, 19% of candidates who scored 89% in the unit opted to resit.

The gradients of the lines for the Chemistry F323 practical unit are shallower than those for written unit F322, indicating that candidates

5. Loess is a non-parametric regression method which avoids imposing a particular model on the data, and is thus well-suited for this exploratory analysis.

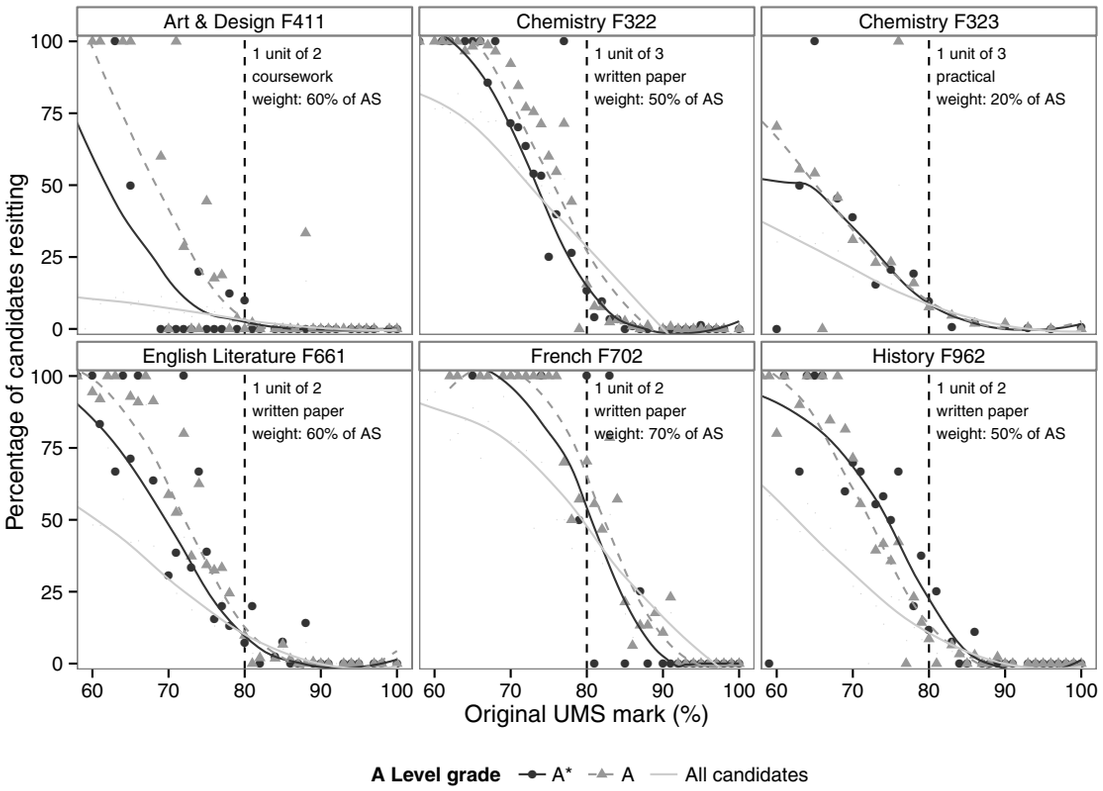


Figure 2: Probability of resitting given original mark, for selected AS units

were not so sensitive to the mark obtained in deciding whether to resit. This is likely to be due to the low weighting of unit F323, meaning that candidates can afford not to resit it and still go on to receive an A*, if they have performed well in the written units.

The Art and Design unit shows a pattern that is different again: overall resitting rates were very low for a particular mark compared to the other units shown here (including Chemistry F323), despite the high weighting of this unit. This may be due to the nature of the assessment (via a portfolio of coursework). However, the best candidates (those awarded A or A* at A level) showed similar resitting patterns as candidates in other units.

Given a particular mark, A* and A candidates were more likely to resit than average; that is, they were less likely to be 'happy' with a particular mark, as one would expect. However, in most of these units, there is little difference between the resitting behaviour shown by A* and A candidates; in fact, in most cases A* candidates were slightly *less* likely to resit, given a particular mark, than those awarded an A. One interpretation could be that A* candidates' resit rates were lower because they were concentrating on A2 units, and did not feel it necessary to shore up their AS results. Thus the lower resit rates could have been more likely to lead to the A*, rather than higher ability leading to lower resit rates.

Estimated grades

One common caveat for many of the foregoing analyses is that they are based on the eventual final grade achieved by a candidate. We can only tell from the results who actually achieved an A*, and not who was aiming for one, which might have a bearing on resitting decisions. Estimated grades are provided by the centre to the awarding body before the final examination session, which arguably reflect the beliefs and aspirations of the candidate and the centre, and thus govern resitting behaviour.

Figure 3 shows whether candidates who achieved (or surpassed) their estimated grade had different resitting patterns from those who did not[6]. There was very little difference in History, English Literature and Art and Design. However, in Chemistry and French, candidates who met their estimated grade had lower resit rates, except at grade E.

This suggests that the more unreliable performance of candidates who have already needed to resit may hinder them from achieving their estimated grade in some subjects.

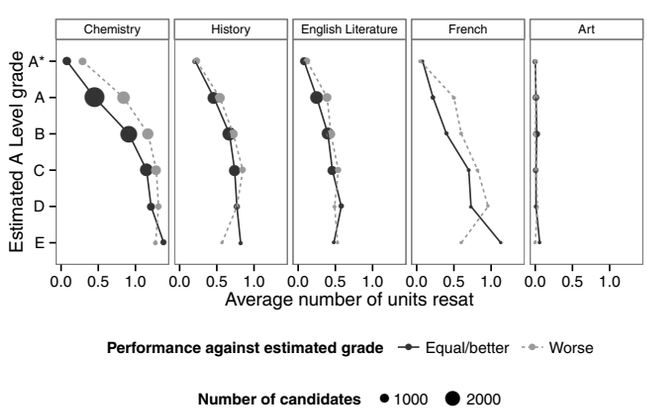


Figure 3: Comparison of resitting and estimated grades

Effect of resitting

Unit mark

Figure 4 shows the average difference between the original mark obtained and the best mark used for certification, across all candidates and units. By definition, this can never be less than zero. The size of each

6. Resitting that occurred in the final session is not included in this calculation; it is restricted to events that had already occurred by the time the centre provided the estimated grades.

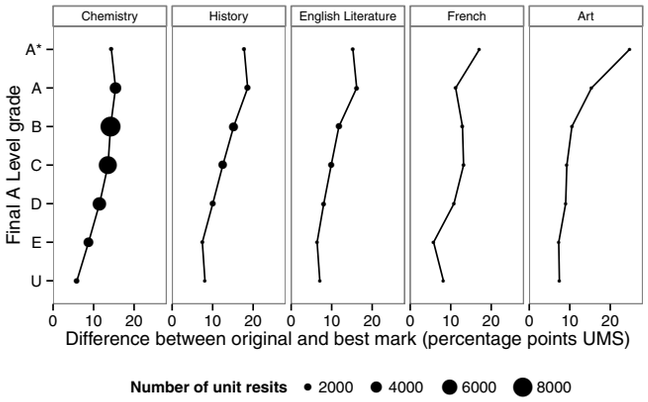


Figure 4: Difference between original and best mark

point represents the underlying number of unit resits used to calculate the average.

In all subjects, the higher ability candidates managed to increase their unit marks the most by resitting. The marks gained by A* and A candidates were similar in most subjects, with the exception of French and Art and Design.

Overall grades

Tables 2 to 6 below show the effect of resitting on overall grade distributions for each of these subjects, by recalculating grades based on the first sittings of each unit (which we have termed the 'original grade') and comparing to those actually awarded. The off-diagonal entries show where candidates improved their certificate grade by resitting units. Under the current rules, it is not possible to end up with a worse grade from resitting than the original (as the highest mark for each unit is used for certification), and hence the bottom left of each table is blank.

In all these subjects, resitting had the effect of increasing the numbers of candidates awarded A*, A and B, at the expense of the lower grades. The largest increases, in both absolute and relative terms, were seen for the A grade, but the increase for the A* grade was modest. The largest increase in the proportion of candidates awarded A* was in History (from 6.7% to 7.5%), whereas Chemistry had the largest increase at A (from 16.3% to 26.6%).

In most subjects, a few candidates managed to increase their overall grade substantially by resitting units: for example, one candidate in each of Chemistry and History would have obtained a D if the first sitting of each unit had been considered, but in fact went on to be awarded an A*.

Table 2: Original & final grade distribution for Chemistry

Grade (with resit)	*Grade (original)*								
	*A**	*A*	*B*	*C*	*D*	*E*	*U*	*Total*	*%*
A*	1514	83	18	1	1	.	.	1617	8.9
A	.	2875	1255	135	20	6	.	4291	23.6
B	.	.	2859	1606	262	33	5	4765	26.3
C	.	.	.	1880	1287	230	17	3414	18.8
D	.	.	.	.	1215	870	99	2184	12.0
E	.	.	.	.	.	935	434	1369	7.5
U	.	.	.	.	.	.	509	509	2.8
Total	1514	2958	4132	3622	2785	2074	1064	18149	
%	8.3	16.3	22.8	20.0	15.3	11.4	5.9		

Table 3: Original & final grade distribution for English Literature

Grade (with resit)	*Grade (original)*								
	*A**	*A*	*B*	*C*	*D*	*E*	*U*	*Total*	*%*
A*	1406	34	28	2	.	.	.	1470	14.5
A	.	1791	360	30	2	.	.	2183	21.6
B	.	.	2398	445	25	.	.	2868	28.3
C	.	.	.	1880	313	20	2	2215	21.9
D	.	.	.	.	952	122	2	1076	10.6
E	.	.	.	.	.	238	20	258	2.5
U	.	.	.	.	.	.	50	50	0.5
Total	1406	1825	2786	2357	1292	380	74	10120	
%	13.9	18.0	27.5	23.3	12.8	3.8	0.7		

Table 4: Original & final grade distribution for History

Grade (with resit)	*Grade (original)*								
	*A**	*A*	*B*	*C*	*D*	*E*	*U*	*Total*	*%*
A*	723	36	49	2	1	.	.	811	7.5
A	.	1694	618	75	7	.	.	2394	22.1
B	.	.	2334	938	88	2	.	3362	31.1
C	.	.	.	1947	659	55	3	2664	24.6
D	.	.	.	.	969	255	6	1230	11.4
E	.	.	.	.	.	271	33	304	2.8
U	.	.	.	.	.	.	45	45	0.4
Total	723	1730	3001	2962	1724	583	87	10810	
%	6.7	16.0	27.8	27.4	15.9	5.4	0.8		

Table 5: Original & final grade distribution for French

Grade (with resit)	*Grade (original)*								
	*A**	*A*	*B*	*C*	*D*	*E*	*U*	*Total*	*%*
A*	117	1	2	.	.	.	.	120	11.0
A	.	289	76	2	.	.	.	367	33.5
B	.	.	191	85	7	.	.	283	25.8
C	.	.	.	112	78	4	.	194	17.7
D	.	.	.	.	62	26	1	89	8.1
E	.	.	.	.	.	27	4	31	2.8
U	.	.	.	.	.	.	11	11	1.0
Total	117	290	269	199	147	57	16	1095	100.0
%	10.7	26.5	24.6	18.2	13.4	5.2	1.5		

Table 6: Original & final grade distribution for Art and Design

Grade (with resit)	*Grade (original)*								
	*A**	*A*	*B*	*C*	*D*	*E*	*U*	*Total*	*%*
A*	540	.	5	.	.	.	.	545	14.7
A	.	758	26	2	.	.	.	786	21.1
B	.	.	1047	36	4	.	.	1087	29.2
C	.	.	.	712	20	.	.	732	19.7
D	.	.	.	.	401	5	1	407	10.9
E	.	.	.	.	.	138	3	141	3.8
U	.	.	.	.	.	.	21	21	0.6
Total	540	758	1078	750	425	143	25	3719	
%	14.5	20.4	29.0	20.2	11.4	3.8	0.7		

Table 7: Comparison of extent of resitting and effect on overall grades

Subject	*All candidates*			*Candidates with an A based on first sittings*		
	Total candidates resitting	*Candidates increasing overall grade*	*As % of resitters*	*Total candidates resitting*	*Candidates increasing overall grade*	*As % of resitters*
Chemistry	13862	6362	45.9	1062	83	7.8
English Literature	3817	1405	36.8	248	34	13.7
History	6437	2827	43.9	394	36	9.1
French	602	286	47.5	72	1	1.4
Art and Design	288	102	35.4	9	0	0.0

In Chemistry, there were far more candidates who moved up to an A* from an A than from a B or lower. The numbers were similar for English Literature, but for History there were more candidates who moved up from a B than from an A. There are two possible reasons for this: first, History students may benefit more from maturational effects, thus being able to make a dramatic improvement towards the end of the course by resitting earlier units. English Literature and History, in particular, require essay writing skills which will develop over the course. Secondly, in all subjects apart from Chemistry most candidates sat both their A2 units for the first time in the final session, so there was less opportunity to resit A2 units. As such, most candidates would only have been able to increase their grade by resitting AS units, meaning that those who achieved an overall A* must have come from a B grade or lower. This is an artefact of the A* rule and the imbalance of opportunity to resit at A2.

Table 7 compares the number of candidates resitting with the number who increased their overall grade by doing so, which might crudely be deemed a success rate for resitting. However, not all resitting will be with the aim of actively increasing the overall grade: some candidates may have resat in an attempt to shore up their grade and reduce the risk of ending up with a lower grade. As shown in Table 7, under half of all resitting was successful in increasing a candidate's overall A level grade, but this rate was much lower for those candidates who would have been awarded an A, on the basis of their first sittings. For the group of all candidates, the rates were remarkably similar across subjects.

Overall mark

Figure 5 shows the total UMS marks obtained by each candidate in both AS and A2, as a percentage of those available (300 marks in Chemistry; 200 in other specifications). The horizontal line shows the threshold of 90% UMS marks at A2, and the sloping solid line shows the A grade boundary, a total of 80% of UMS marks across both AS and A2. To obtain an A* it is necessary to meet both these conditions, so candidates awarded an A* lie in the top right region of the graph. Three scatter plots are presented for each subject. The first two plots show only those candidates who have resat one or more units, and compare the UMS totals derived from the best marks scored in each unit and used for certification ('After resitting') with the totals that would have been obtained had the marks from the first sitting of each unit been used ('Original mark'). The third scatter plot shows the final marks used for certification by all candidates (including those that did not resit any units). The points corresponding to each candidate are shaded according to the actual A level grade obtained. For clarity, only those candidates awarded an A*, A or B in the A level are shown on the plots.

The plots give a visual overview of the effect of resitting: points lying outside their 'zone' in the first plot for each subject indicate that the candidate has increased their grade through resitting. Of particular interest are the points above the 90% horizontal line, but to the left of the A* region in the top right hand corner. In all subjects there are points visible in this area in the first plot (corresponding to the marks awarded on the first sittings of each unit), but most candidates were able to increase their AS marks. Thus the second plot for each subject shows very few candidates in this area. In the final plot it can be seen that in Art and Design, English Literature and History some candidates ended up certificating with a B grade because of this.

A further finding of interest is that the nature and strength of the relationship between AS and A2 marks differed between subjects: for French, AS marks were generally higher than A2 across all grades, whereas for Chemistry this tendency was reduced for candidates of the highest ability, and for English Literature AS marks were more similar to A2 marks. Chemistry also exhibited a high degree of correlation (r=0.85) between AS and A2 marks, and as a result the cluster of candidates on the scatterplot representing all candidates was much tighter than English Literature and History, for example, which had lower correlations (0.76 and 0.65 respectively). This has implications for the possibility of being awarded a B while gaining over 90% of the A2 marks.

Discussion

Resit rates for candidates who were awarded an A* at A level were approximately half the average, and markedly less than those for A candidates (which were around 80% of average). The finding that more able candidates resit less often is consistent with previous research (Gill, 2010; CERP, 2012) and has a number of possible interpretations: performance of excellent candidates could be intrinsically more reliable, or an isolated lower mark may be of less concern to better candidates as they know they will still end up with a good grade. Alternatively, it may be harder to move up a grade through resitting if performance at these levels is qualitatively different. The starred grade was perhaps also not seen as so important if students were on track to go to their chosen university, as those offers that required an A* typically did not specify the subject, so candidates just missing out on an A* grade in one subject may have achieved one in another.

French and Chemistry had much lower resit rates for A* candidates than the average for all candidates. The resit rate for French was particularly low, perhaps indicating that A* candidates achieved a particular degree of mastery of this subject which meant resitting was rarely required; the corresponding rate for A candidates was very close to the overall trend. In addition, French is one of the subjects that is most often dropped after AS (Gill, 2009) and this has implications for the

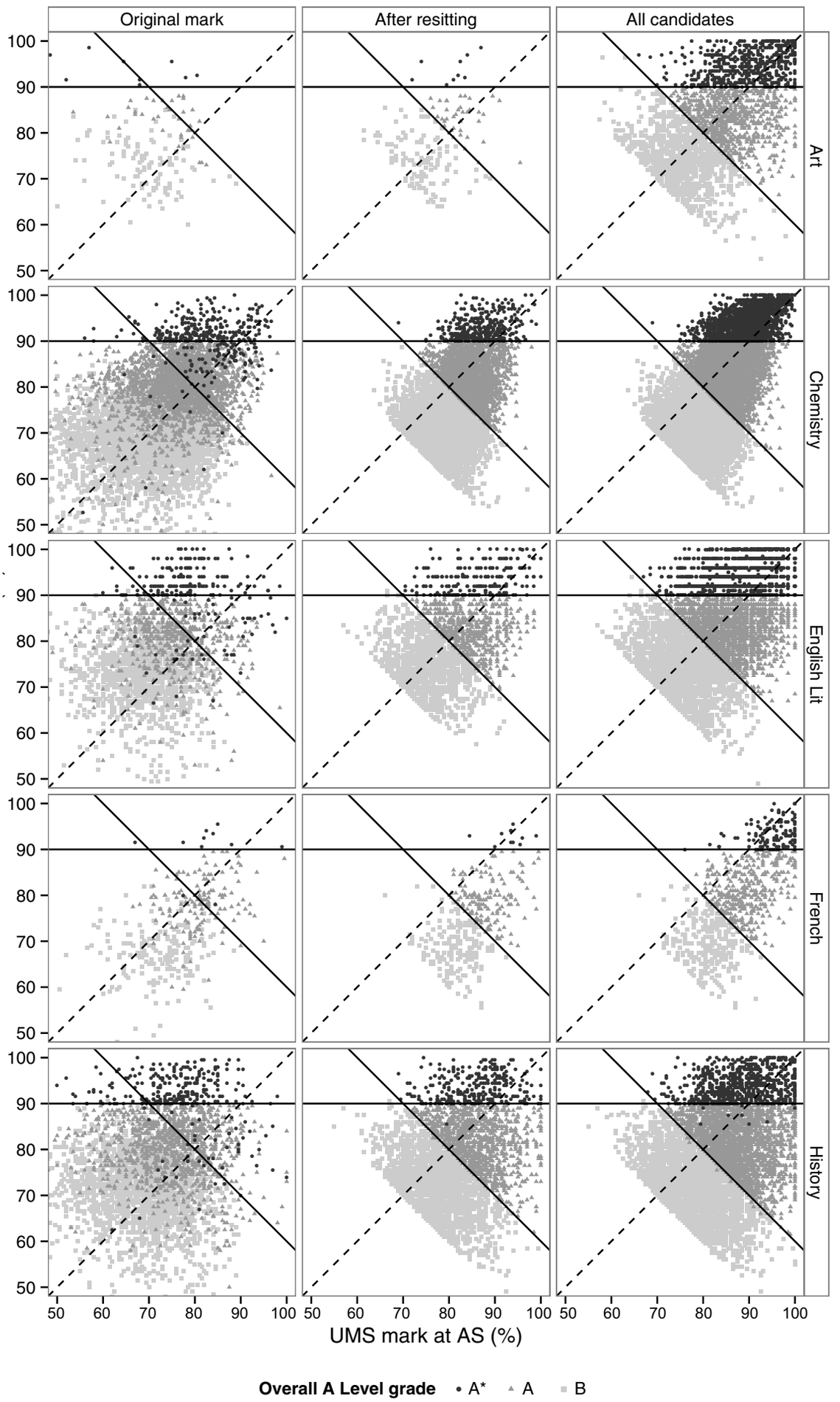


Figure 5: Plots of AS and A2 marks for resitters and all candidates

cohorts taking each half of the A level, which might have affected awarding.

As with the rest of the cohort, A* candidates were more likely to resit AS units than A2 units, which is probably chiefly due to opportunity: many A2 units are taken in the final June session which makes resitting impractical.

We examined the resitting decisions given a particular mark on the first sitting of the unit, and found that A* candidates were in general little different from A candidates in their decision to resit AS units given a unit mark. One finding of interest was that some AS units (such as History F962 and French F702) were resat by candidates after achieving marks of over 80% in the first sitting. Attempting to score a higher mark than this would have little effect on their chances of achieving an A* (unless they performed poorly on another AS unit and were in danger of being awarded a B). It is possible that these candidates misunderstood the A* criteria, but it is most likely that they were seeking insurance against the harder A2 units to ensure they got an A. This was predicted by QCA (2007) along with the suggestion that some excellent candidates would resit in pursuit of a few extra marks to satisfy higher education admissions requirements. Neither of these concerns relates directly to the design of the A*, but rather to the availability of detailed mark scores. A remark made by one of Poon Scott's (2012) interviewees, a student at an independent college, is particularly telling: "My teachers said there's nothing really to lose if you resit because you still have your old grade. Even when you have an A, there's always a better A." (p. 441). In History unit F962, those candidates who eventually went on to get an A* were more likely to have done this than those who got an A, perhaps showing evidence of perfectionism rather than aversion to risk. This is a peculiar feature of modularisation and is due to the time delay between the assessments – in a purely linear A level there would be no opportunity or need to do this.

The net effect of resitting was to increase candidates' unit marks, as expected given the contribution of regression to the mean, the rule that the best unit performance is used for certification, and maturation and further learning between sittings. For most subjects, candidates awarded the A* showed the highest average increase in unit marks between sittings (comparing the first to the best sittings), which is consistent with previous findings by Al-Bayatti and Jones (2003) that candidates with higher mean GCSE scores exhibited the highest increase in marks between sittings of AS units. One explanation for this result could be that better candidates tended to resit because they had had a 'bad day' and wished to correct their performance, rather than taking advantage of the inherent variability of the assessment process to try to gain a few marks and therefore push their mark over a grade boundary. Alternatively, maturation may have a different effect at different points on the ability scale. However, this result should be treated with some caution as large increases in unit marks will tend to have the effect that a candidate's overall grade will increase, so such candidates are less likely to have obtained a U, for example.

Across the whole A level, resitting did not have a large impact on the percentage of candidates awarded an A*: the increase in all subjects considered here was under one percentage point. The proportion of candidates moving up a grade through resitting (throughout the grade distribution) was strongly associated with the proportion of candidates resitting in each subject: approximately half of resitters increased their overall grade by doing so, with this ratio remarkably consistent between subjects. This is somewhat lower than perceptions from teachers: in de Waal's (2009) survey, most teachers thought that more than half of their students (that is, not just as a proportion of those that had resat) had improved their overall grade through resitting, and 30% of teachers thought that 80% of their students had improved their overall grade. While our study does not cover all subjects, is based on new specifications as opposed to the old six-unit versions and is restricted to OCR candidates, there does seem to be a gulf between these perceptions and the statistics.

In History, a greater number of A* candidates would have been awarded a B than an A if they had not resat any units. This suggests that resitting AS units is implicitly being rewarded by the A* criteria due to the nature of the relationship between AS and A2. In History one might expect maturation and further learning to make resitting AS units particularly successful as writing skills are honed and the student gains more appreciation of broader context, but a large element of this may simply be down to opportunity (as observed by QCA, 2003c). The majority of candidates sit both A2 units for the first time in the final June session and there is thus little opportunity to resit them without affecting future plans for employment or education. In turn this means that few candidates are able to move from an A to an A* by resitting, as this would necessitate resitting A2 units, and therefore most of the resitters come from a grade B. The contribution made to grades and resitting by sheer opportunity is potentially of concern as this varies across specifications, due to the patterns of entry for unit exams which are in turn driven by the numbers and weighting of units. Additionally, the imbalance of opportunity may be inequitable between different types of schools and colleges, depending on when students are entered for unit assessments for the first time.

In some subjects, candidates who were forecast an A* and went on to achieve it resat fewer units in the A level than those who did not, so resitting earlier in the A level (perhaps indicating unreliability of performance) may be a predictor for poorer or unreliable performance in the final session. As well as any inherent character traits this might reveal, resitting could also increase the overall examination burden in the final session and distract the candidate from their A2 units. This may justify use of information on resits at AS as additional information for higher education admissions.

Following the recent consultation on A level reform, it was announced (Ofqual, 2012) that no January exam sessions will be available from 2014 onwards (for the remainder of the life of the modular A levels). This will mean that students will have only one opportunity to resit AS units, and A2 units will typically be taken in the final June session. In practice the A2 units are largely already taken in the final session for History and English Literature, and in Art and Design only summer assessment is available at present. However, there will be an effect on patterns of entry and resitting in Chemistry. This study has shown that the highest achieving candidates will be less affected by this than other candidates, as they resit less, although the small number of candidates who do resit due to an anomalous performance may be disadvantaged. From 2015, a subsequent reform of A levels looks set to remove the AS in its current form, and move to a fully linear qualification, in which resitting will be much less practical. The patterns of resitting observed in this study by students achieving different grades suggest that the highest achieving students are overall less likely to be affected by the reform than less able students.

Conclusion

This study investigated resitting in high achieving A* candidates at A level compared to other candidates by looking at both the extent and effect of resitting. We have found evidence that resitting at A level is not as prevalent or as effective as popularly supposed. Overall, high achieving candidates were less likely to resit units than other candidates. However, the increase in marks after resitting was higher for A* candidates than for other candidates, suggesting that A* candidates might be more likely to resit due to an anomalous performance. Since more highly achieving candidates are less likely to resit units, information on a student's resits could be a good predictor for A level performance, and provide useful information for higher education admissions.

References

Al-Bayatti, M.F., & Jones, B. (2003). *Statistical study of the differences in candidates' results between first and second attempts in some GCE AS units*. Paper presented at the British Educational Research Association Annual Conference, September 11–13, in Edinburgh, UK. Available online at: http://www.leeds.ac.uk/educol/documents/00003197.htm (Accessed 28 November 2012).

CERP (2012). *What is the impact of resitting at A-level?* AQA Centre for Education Research and Policy, Manchester. Available online at: http://cerp.aqa.org.uk/sites/default/files/pdf_upload/CERP_IP_resitting_14072012.pdf (Accessed 28 November 2012).

Dearing, R. (1996). *Review of qualifications for 16–19 year olds*. London: SCAA.

de Waal, A. (2009). *Straight As? A-level teachers' views on today's A-levels*. London: Civitas. Available online at: http://www.civitas.org.uk/pdf/straight_a's.pdf (Accessed 18 January 2013).

Gill, T. (2009). *Uptake of GCE AS level subjects in England 2001–2006*. Statistics Report Series No.8. Cambridge: Cambridge Assessment.

Gill, T. (2010). *Patterns and impact of re-sits in unitised GCSE and A-level specifications*. Research Report. Cambridge: Cambridge Assessment.

Gill, T. & Suto, I. (2012). *Students' and teachers' views and experiences of A level module re-sits*. Paper presented at British Educational Research Association conference, Manchester, September 2012.

Gray, E. A. (2002). *The comparability between modular and non-modular examinations at GCE Advanced level (PhD thesis)*. Institute of Education, University of London. (Accessed from EThoS Persistent ID: uk.bl.ethos.251901).

Higton, J., Noble, J., Pope, S., Boal, N., Ginnis, S., Donaldson, R. & Greevy, H. (2012). *Fit for purpose? The view of the higher education sector, teachers and employers on the suitability of A levels*. Ofqual/12/5145. Available online at: http://www.ofqual.gov.uk/files/2012-04-03-fit-for-purpose-a-levels.pdf (Accessed 18 January 2013).

MacLeod, D. (2007, September 21). No limit on A-level resits, watchdog rules. *The Guardian*. Available online at: http://www.guardian.co.uk/education/2007/sep/21/schools.alevels (Accessed 25 January 2013).

Matthews, A., & Pepper, D. (2007). *Evaluation of participation in GCE mathematics*. QCA/07/3388. London: QCA. Available online at: http://www2.ofqual.gov.uk/downloads/category/106-gq-monitoring (Accessed 25 January 2013).

Ofqual (2010). *Open letter to secondary schools and colleges about the new A* grade at A level*. Available online at: http://www2.ofqual.gov.uk/files/2010-02-11-open-letter-a-star-grade.pdf (Accessed 18 January 2013).

Ofqual (2012). *Ofqual announces changes to A levels*. Press release, 9 November 2012. Available online at: http://www.ofqual.gov.uk/news/ofqual-announces-changes-to-a-levels/ (Accessed 18 January 2013).

Poon Scott, E. (2012). Short-term gain at long-term cost? How resit policy can affect student learning. *Assessment in Education: Principles, Policy & Practice*, **19**, 4, 431–449. doi:10.1080/0969594X.2012.714741

Qualifications and Curriculum Authority (QCA), (2003a, 27 January). *Review of A level awarding in 2002*. Available online at: http://web.archive.org/web/20030625040443/http://www.qca.org.uk/nq/framework/review_a_level_grading.asp (Accessed 24 January 2013).

Qualifications and Curriculum Authority (QCA), (2003b, 3 October). *Revision of A-level resit regulations*. Available online at: http://wayback.archive.org/web/20080915101555/http://www.qca.org.uk/qca_9696.aspx (Accessed 24 January 2013).

Qualifications and Curriculum Authority (QCA). (2003c, December). *Curriculum 2000 Review: Report on Phase 3*. Available online at: http://dera.ioe.ac.uk/id/eprint/10427 (Accessed 24 January 2013).

Qualifications and Curriculum Authority (QCA), (2007). *A-level resitting: summary of research findings*. Available online at: http://www2.ofqual.gov.uk/files/qca-07-3387-Resit-report.pdf (Accessed 18 January 2013).

Smith, G., & Smith, J. (2005). Regression to the mean in average test scores. *Educational Assessment*, **10**, 4, 377–399. doi:10.1207/s15326977ea1004_4

Tomlinson, M. (2002). *Inquiry into A level standards: final report*. London: Department for Education and Skills.

Wheadon, C. (2010). *The impact of re-sitting on the fairness of grading*. AQA Centre for Education Research and Policy, Manchester. Available online at: http://cerp.aqa.org.uk/research-library/impact-re-sitting-fairness-grading (Accessed 18 January 2013).

Williams, D. A. (2009). *What has been the impact of re-sitting AS-Level examinations in Economics and Business Studies on students at a boys' independent school in the West Midlands?* (PhD thesis). University of Warwick. Available online at: http://wrap.warwick.ac.uk/2803/ (Accessed 18 January 2013).

Comparing difficulty of GCSE tiered examinations using common questions

Vikas Dhawan and Frances Wilson Research Division

Introduction

Tiering is a test design followed in the UK for some GCSE examinations whereby it is intended to develop tests at different difficulty levels (and with different available grades). Teachers or schools then decide what the most appropriate tier is for their pupils. In such a *differentiated* assessment the higher proficiency candidates are allocated to the more difficult 'higher tier', whereas those towards the lower end of the proficiency scale are allocated to the easier 'foundation tier'. The foundation tier covers grades G to C and the more difficult higher tier covers grades D to A*, with grade E often allowed for those candidates who just miss grade D. The overlapping grades in the two tiers, C and D, are intended to represent the same level of performance, irrespective of the tier on which they may be achieved. Table 1 shows the grades available on the foundation and higher tier components of a tiered GCSE unit.

Table 1: Grades available on GCSE tiered components

				Overlapping grades						
Higher tier grade	A*	A	B	**C**	**D**	*E*	Ungraded			
Foundation tier grade				**C**	**D**	E	F		G	Ungraded

The process of setting grade boundaries (the minimum mark required to attain a grade) for each examination for each session is called 'awarding'. It is based on the procedures laid down in the Code of Practice (Ofqual, 2011a) issued by the examination regulator – Office of Qualifications and Examinations Regulation (Ofqual) – which states that the purpose of awarding is "to ensure that standards are maintained in each subject examined from year to year…". Subject matter experts compare candidate scripts (or coursework, if applicable) at different performance levels and judge them to be worthy of specific grades. The grade boundaries are decided by the experts based on the candidate performance and various sources of statistical evidence (such as teachers' forecast grades, performance of the cohort in the previous sessions, etc.). The complete list of potential sources of evidence which can be used for awarding is given in Ofqual (ibid.). The grade boundaries which are decided by the experts by using these sources of evidence are called 'key boundaries'. Not all grade boundaries are obtained by this 'judgemental' process. The rest are calculated arithmetically to lie between the key boundaries. For the GCSE tiered examinations used in this study, the key boundaries are A, C and F (as well as D on the higher tier only[1]).

In tiered examinations, a comparison of the performance at the overlapping grades can be used to maintain standards between the two tiers. Usually the performance at the common grade C (the highest possible grade on the foundation tier) is used to achieve this objective. This is done by developing some items which are common to both question papers. In this study we define 'common items' as those items[2] which have exactly the same structure, format and wording across the two tiers. Usually the mark scheme for common items should be identical across tiers. Here we distinguish between common items and 'similar' items. We define similar items as those items which test the same content across tiers, but use different wording or question structure. The rest of the questions are unique to each paper and have been referred to as 'non-common' questions in this study. It is intended that the non-common questions should on average be easier than the common questions in the foundation tier and should in general be more difficult than the common questions in the higher tier so that the foundation tier is easier than the higher tier. Such a question paper design is intended to provide more support to the foundation tier candidates and more stretch to the higher tier candidates.

If the questions were not functioning as intended, the foundation tier might be more difficult or the higher tier easier than they should be. This essentially is an issue related to test construction and could lead to unexpected differences in the C boundaries between the two tiers. It is normally expected that the C boundary on the foundation tier should be set at a higher proportion of the maximum numeric mark (paper total) than on the corresponding higher tier, because due to the differences in the difficulty of the tiered papers, pupils at this level of achievement (i.e. grade C) should have to get a greater proportion of marks on the foundation tier to attain the same grade than on the relatively more difficult higher tier. If the C boundaries did not function according to this criterion, then it would suggest that the questions on each tier had not been targeted effectively – the foundation tier questions might be too difficult, or higher tier questions too easy. This conclusion would be relevant even if common items were not used. A negative or a very small difference between the C boundaries could indicate that one or both of the question papers might not have been at the target difficulty. Under these circumstances, there is a risk that the grades received by candidates might not be an appropriate reflection of the level of their understanding or proficiency in the subject area. Along with test construction, another reason why the boundaries could be set at the unexpected place is related to awarding. If the grade boundaries were not set appropriately, the difference between the C boundaries could be negative or very small. The use of common questions allows us to investigate further what the 'real' reason might be.

In this study, we investigated the difficulty of the common questions between tiered components to gather evidence of whether the tiered question papers were functioning as expected or not. We used data from the awarding body OCR. We also explored ways in which the analyses could feed into the process of writing questions for tiered examinations and thereby help in improving the current practice of producing such question papers.

1. True for the data used for this study. From June 2012, however, grade D on the GCSE higher tier is now calculated arithmetically.
2. Sub-parts of a question.

For a more theoretical understanding of how tiered examinations work, see Good and Cresswell (1988a, 1988b, 1988c). Wheadon and Béguin (2010) also give a useful discussion on improving standard setting on tiered tests using Item Response Theory (IRT) models.

Current practice in producing common questions and tiered papers

Tiered assessments have been a common feature of GCSE assessment since the introduction of GCSEs in 1988. Towards the end of the 1990s, the number of tiers used to differentiate between candidates was reduced from three to two in most subjects. The number of tiers to be used in GCSE assessments is regulated by Ofqual. In 2004 the Qualifications and Curriculum Authority (QCA) – Ofqual's predecessor – specified that "the assessment arrangements for GCSEs must … include question papers targeted at two tiers of grades, A*-D and C-G, unless subject criteria or the regulatory authorities indicate otherwise" (QCA, 2004, p.27). However, Ofqual's Code of Practice (2011a) does not specify whether GCSE assessment should be tiered or not. Currently, the Subject Criteria, which are published by Ofqual on a subject by subject basis, specify whether a given subject must have tiered examinations. To the best of our knowledge, no formal motivation for the decision to use tiered examinations or not in a subject is available. It is frequently informally suggested that subjects which have tiered examinations are those such as Science or Mathematics, where the questions targeted at the top grades would be inaccessible to less able candidates, and would thus provide a demotivating assessment experience for these candidates. However, it is unclear, for example, why Latin is tiered, but Classical Greek is not (Ofqual 2011b), or why History is tiered in Northern Ireland, but not in England and Wales (Ofqual, 2012a).

Tiered question papers aim to differentiate candidates of different abilities, while still allowing for comparability between the awarded grades where the papers overlap (grades C and D). There are often differences between foundation and higher tier papers with respect to the style and format of tasks. Foundation tier papers frequently use tasks which are more structured, and use less complex vocabulary and sentence construction. However, common items should be suitable for use in both foundation and higher tiers. Within OCR, there are few formal guidelines for setting common items for those subjects which have tiered examination papers. Where such guidelines exist, they do not typically extend beyond specifying the need to target common items at the overlapping grades C and D. For some OCR question papers, for example, 2359 (ICT) or A353 (Classical Civilisation), the common questions are set in a block of questions which differ in format from the questions specific to the foundation tier (typically objective) and higher tier (typically extended answer). Where common questions form a block of questions, they typically occur towards the end of the foundation tier paper, but at the beginning of the higher tier paper, consistent with the more general approach of putting the most difficult exam questions towards the end of a paper.

Despite the general lack of literature relating to the current practice in setting common items, Ofqual does consider the use of common items in some papers. In a review of standards in GCSE English between 2005 and 2009 (Ofqual, 2011c), it was noted that the tasks common to both foundation and higher tiers may provide more scaffolding than is appropriate for the higher tier. In contrast, a review of GCSE Business Studies (QCA, 2005a) commented that where similar questions were used across tiers (in OCR and AQA 2003 examinations), the question format was more similar to that used in the higher tier papers, and was rather demanding for the foundation tier candidates. However, if different guidelines for question formats are set for different tiers, it seems inevitable that the format of common tasks must compromise tier specific guidelines to some extent.

Ofqual also conducted a series of recent reviews of GCSE standards across time in a limited set of subjects (Ofqual, 2012b). A minority of papers in this series comment favourably on the use of common questions. For example, a review of Biology GCSE standards from 1999–2003 (QCA, 2005b) notes that the use of common questions allows comparison of standards for candidates awarded grade C across tiers. Additionally, a review of Chemistry GCSE standards (QCA, 2005c) in 1998 and 2003 recommended that where common or similar questions are used, identical wording should be used to allow direct comparison across tiers, even if this means that foundation tier wording is used on higher tier papers. It is notable that this is contrary to the recommendations for English (Ofqual 2011c), which criticised the use of foundation tier question wording in a higher tier paper.

Method

A list of 81 pairs of foundation-higher tier assessments (referred to as 'component-pairs' in this study) were obtained from the June 2011 session examinations. The C boundary raw mark of each component was calculated as a percentage of its paper total and a difference (foundation – higher) between the percentages of each component-pair was used to select a potential group of component-pairs. A positive difference here would indicate the expected situation – that the C boundary on the foundation tier as a proportion of its paper total was higher than that on the higher tier. On the other hand, a negative (or a very small positive) difference might suggest that issues related to test construction and/or awarding need to be investigated. The difference between the C boundaries in each pair was also compared against the 'target' or expected difference set by OCR at 45 percentage points – 85% of the foundation paper total and 40% of the higher paper total (Dhawan, 2012).

Figure 1 shows the distribution of differences (foundation – higher) in C boundaries as a proportion of paper total between the two tiers for 81 pairs for the June 2011 session. The graph also gives a few summary values of the differences. A vertical line at 0.0 on the x-axis identifies the point where the C boundary between the two tiers was exactly the same. Another vertical line at 45.0 identifies the target difference in C boundaries between the tiers.

Figure 1 shows that the mean difference in C boundaries in the component-pairs was 23.2 percentage points with a standard deviation of 6.9. The median of the differences was 23.8 and the minimum and maximum difference was 8.3 and 42.5 respectively. As is evident from the figure, no negative difference values in the C boundaries were observed. However, there were a few components with a very low difference in the C boundaries which could flag up some possible concerns in the design of the question papers. There were hardly any pairs which were close to the target difference of 45 percentage points. Table 2 gives the summary statistics of C boundaries as a percentage of paper total for all the components. The table suggests the C boundaries at the higher tier were, on average, close to

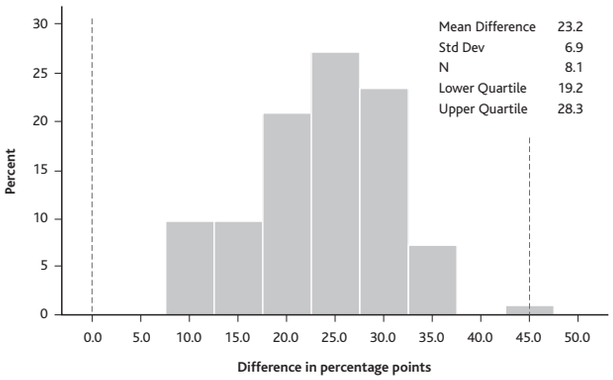


Figure 1: Differences in grade C boundaries (as a percentage of maximum mark) between foundation and higher tiers, June 2011.

the target (40% of raw marks) whereas those at the foundation tier were, on average, lower than the target (85% of raw marks). The lower than targeted C boundaries at the foundation tier indicates why there were hardly any component-pairs near the target difference.

Table 2: Statistics of C boundaries as a percentage of paper total

Tier	**Mean**	StdDev	N	Min	Max	Q1	**Median**	Q3
Foundation	**65.5**	11.1	81	43.6	92.5	57.5	**63.9**	70.4
Higher	**42.3**	12.6	81	18.3	68.0	31.0	**40.0**	52.0

The results given here for the June 2011 session (and those for the June 2009 and June 2010 sessions given in Dhawan, ibid.) suggest that OCR might have set itself a demanding target of achieving a 45 percentage point difference between the C boundaries.

A final list of six component-pairs selected based on the level of difference between the C boundaries of the two tiers and review of question papers and mark schemes is given in Table 3. Note that we have given generic labels to the components. Two component-pairs with the largest unexpected difference, two with the most commonly observed difference between C boundaries and two with a large positive difference were classified respectively as:

a. **Low group** – a difference of less than 17 percentage points;

b. **Median group** – around the average difference of all the pairs; and

c. **High group** – around the target difference of 45 percentage points.

The table also gives the difference between the C boundaries in each pair and the percentage of common items in the paper.

The functioning of items between each component-pair was investigated using Rasch analysis. The Rasch method expresses the estimates of item difficulty and candidate ability on the same scale, called logit or log odds unit scale (Bond and Fox, 2007). This method produces estimates of relative item difficulty which are independent of the ability of the cohort and estimates of candidate ability which are independent of the difficulty of the items. First, we compared the relative ordering of the difficulty of the common items in the tiers. Secondly, we used the common items to equate the two tiers in each pair and compared item difficulty with cohort ability distribution. The common items were used to equate the two tests by applying what is known as the one-step or concurrent method (Hanson and Béguin, 2002; Morrison and Fitzpatrick, 1992). In this method, the student responses from both the tests to be equated are combined in a single dataset and the calibration of the tests is done simultaneously.

Kolen and Brennan (2004, p.271) recommend that the common items should be at least 20% of the length of the total test for equating to be adequate in practice. This is on the assumption that the examinee groups are not very different. However, in the context of tiered exams we are dealing with rather different groups and, as Klein and Kolen (1985) (cited in Cook and Petersen, 1987) demonstrated, "when examinee groups are different the proportion of items common to the tests becomes more important". Table 3 shows that the percentage of common items appeared acceptable for all the component pairs.

We then conducted a qualitative review of how common items relate to non-common items within a pair of question papers, and examined how similar questions, which test the same or similar content, varied across the tiers.

Results

Comparing relative difficulty of the common items

The item difficulty values from Rasch analysis were compared for the common items in each pair. If the foundation and higher tiers are assessing the same trait, differing only in overall difficulty, then the common items should have the same *relative* difficulty in both. Data from both the tiers in each pair were analysed separately and the difficulty values of the common items were plotted against each other[3].

3. The separate analyses fix the origin of each scale at the mean item difficulty (i.e. including common and non-common items) on each tier. Therefore the common items will have a different mean difficulty in each tier. The two scales are aligned by 'shifting' the values from one of the tiers by an amount equal to the difference in mean difficulty of the common items (see Wright & Stone, 1979, pp.112–118).

Table 3: Difference in C boundaries as a proportion of paper total between the tiers

Group	*Component-pair label*	*Subject*	*C boundary*		*Paper Total*		*C boundary/ Paper Total %*		*Difference in % pts*	*% of common items*	
			F	*H*	*F*	*H*	*F*	*H*	*F - H*	*F*	*H*
Low	L1	Biology	24	19	55	55	43.6	34.5	9.1	27.6	28.6
Low	L2	Additional Applied Science	24	18	36	36	66.7	50.0	16.7	34.8	38.1
Median	M1	Applied Science	37	23	60	60	61.7	38.3	23.3	41.2	41.2
Median	M2	Physics	31	17	60	60	51.7	28.3	23.3	27.1	28.9
High	H1	Mathematics	35	17	60	60	58.3	28.3	30.0	16.7	19.2
High	H2	ICT	32	11	60	60	53.3	18.3	35.0	50.0	61.5

F=Foundation tier H=Higher tier

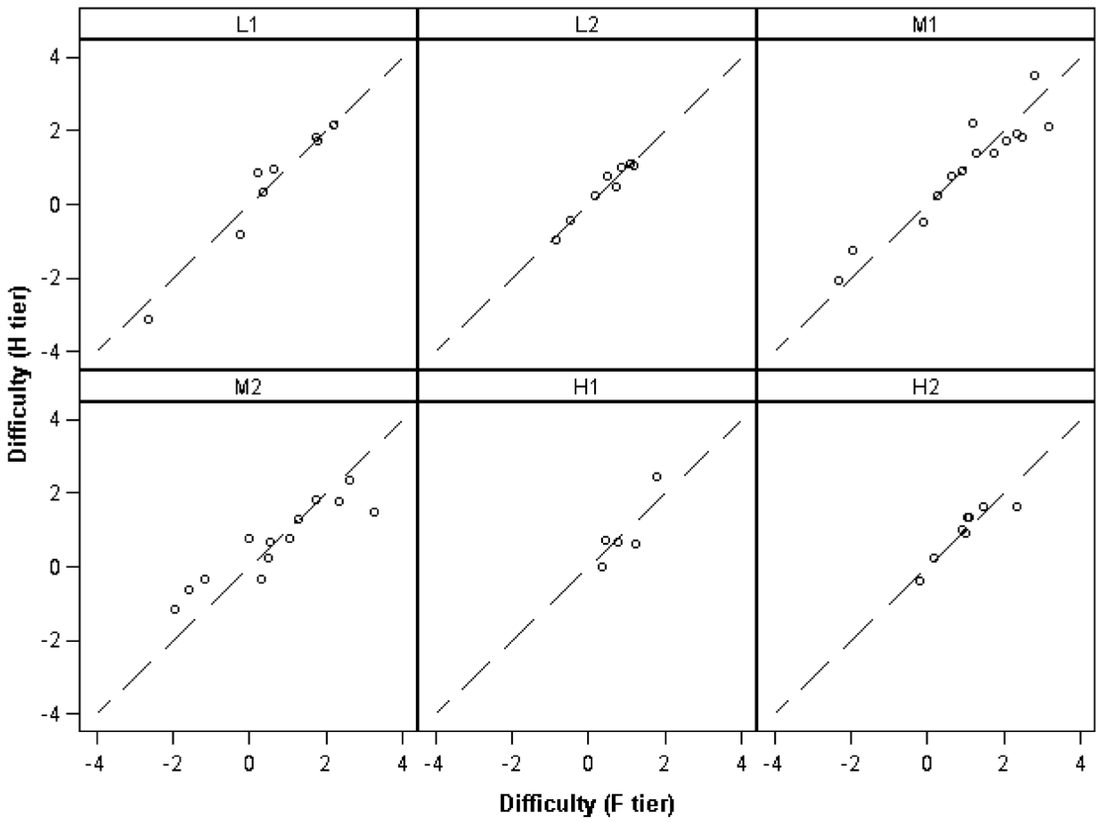


Figure 2: Comparison of Rasch difficulty of common items – Foundation and Higher tier

The results are shown in Figure 2. The x-axis shows the Rasch difficulty on the logit scale in the foundation tier and the corresponding values in the higher tier are given on the y-axis. The items towards the negative end of the scale (-4) indicate easier items, whereas those towards the positive end indicate more difficult items. The line in the middle of each plot is an identity line. Items that fall on this line had the same difficulty values across the tiers. The items below this line were relatively easier on the higher tier, whereas those above the line were relatively easier on the foundation tier.

Figure 2 shows that most of the common items across the six component-pairs were either on or close to the identity line and therefore were of similar difficulty between tiers. There were a few common items, particularly in the Median group (M1 and M2), which did not appear to have similar difficulty and therefore might not have been adequately functioning as common items. From this figure, it appears that the common items in the pairs that were classified into the Low group (L1 and L2) and the High group (H1 and H2) were more or less of similar difficulty. However, it might be due to the fact that the pairs in the Median group had a higher number of common items, some of which did not function as intended.

Overall, it appears that the common items in almost all the pairs had the same relative difficulty on the foundation tier and the higher tier, suggesting that it is reasonable to use common items equating to link scores across the tiers.

Item difficulty and cohort ability

The results from Rasch common item equating for each component-pair are given in Figure 3. The lower part of the graph for each pair shows the estimates of item difficulty after equating. The items towards the left hand side on the x-axis are the easier items and become increasingly difficult towards the right hand side. The items have been identified as common (shown as dots), non-common in the foundation tier (triangles) and non-common in the higher tier (squares). The item estimates are shown here after equating; therefore the common items appear at the same position for both the tiers. The upper part of the graph shows the percentage distribution of ability estimates of pupils on both the tiers. The graph also gives the number of pupils for each tier. Pupils with lower proficiency in this test are shown towards the left hand side of the x-axis and those with higher proficiency are towards the right hand side.

Figure 3 shows that in some components such as L2 and H2 there was hardly any difference in the ability of higher tier and lower tier candidates which suggests that the use of tiering is redundant in these assessments. The figure also shows that the non-common items in some components such as L1 and L2 were very similar in difficulty contrary to the expectation. This effect tends to improve with the increase in the difference in the C boundary between the tiers and the components H1 and H2 have a more clear distinction between the items in the two tiers. H2 gives the best example in this study of the relation between the common and non-common items in which the common items were the easier ones in the higher tier and the more difficult ones in the foundation tier.

The distribution of common items with respect to non-common items is partially dependent on the distribution of items targeted at specific grades. While not all specification grids[4] for the papers analysed in this study give specific grade-targeting information, the specification grids for four of the six papers (H2, M1, M2 and L1) show that common items were indeed targeted at grades C and D, as expected. However, in four of these papers (M1, M2, L1 and L2) the higher tier papers also included

4. A specification grid of a question paper gives a mapping table of items to their target grades.

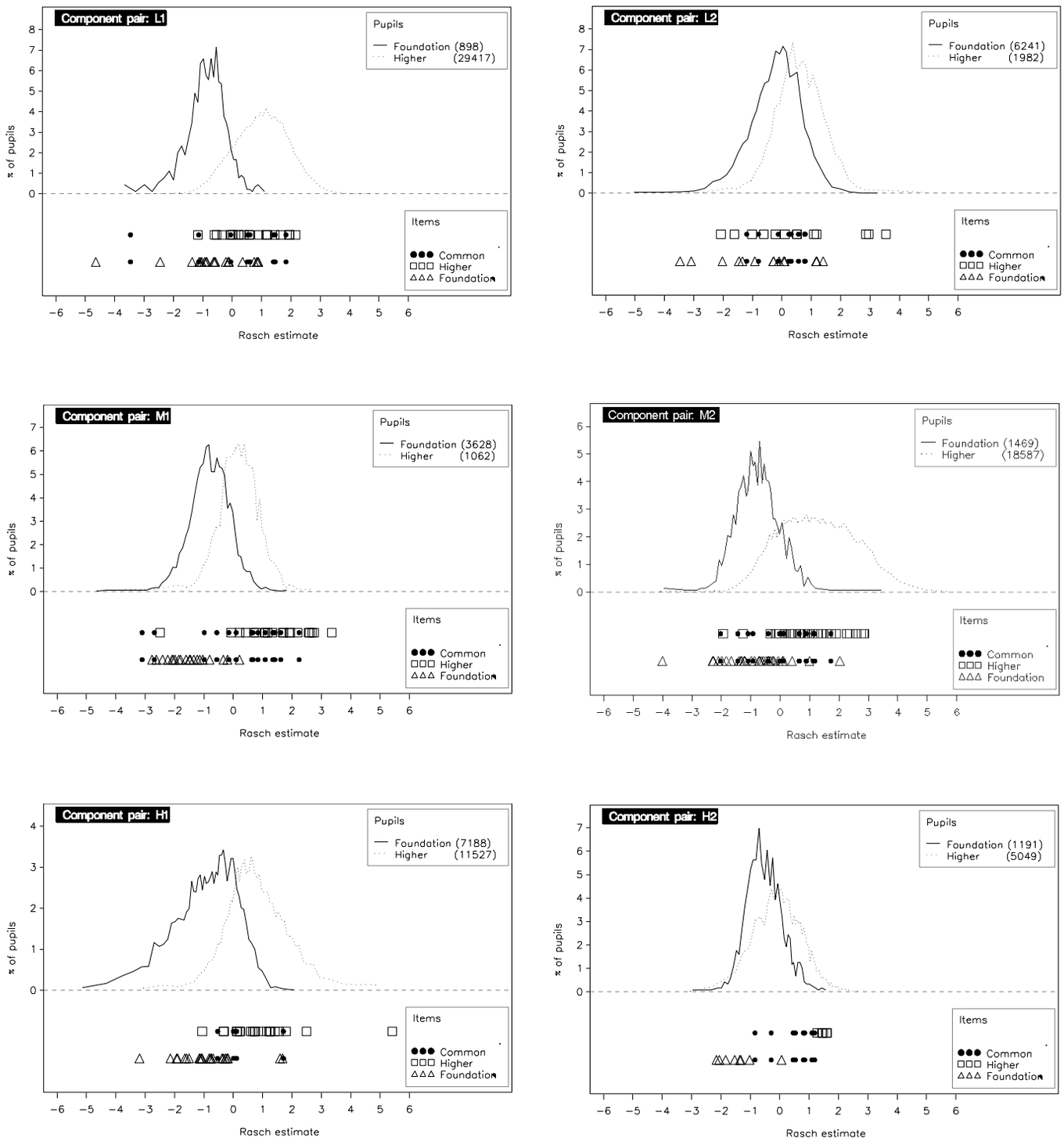


Figure 3: Rasch common item equating, foundation and higher tier

non-common questions which were targeted at grades C and D, whereas the foundation tier paper did not. A closer examination of the non-common items targeted at grades C and D indicates that these items were not as easy as expected for questions at these grade levels in the higher tier. None of the four papers which had non-common items targeted at grades C and D had a high difference between grade C boundaries. For paper H2 (ICT), which had a greater difference between C boundaries across tiers, only common items were targeted at the C and D grades. It seems plausible, therefore, that including non-common items targeted at overlapping grades in the higher tier may have contributed to raising the grade C boundary in the higher tier.

Qualitative review of common items

The question papers surveyed showed different strategies for integrating common and non-common items. Papers H2 (ICT) and M1 (Applied Science), presented common items within a block of questions, which was at the beginning of the higher tier paper and the end of the foundation tier paper, reflecting the fact that the items which are more challenging are typically presented at the end of the paper. H1 (Mathematics) and L1 (Biology) followed a similar pattern, with some variation. Separating common items into one block of questions helps to allow comparability by reducing context effects from other question

parts. However, it is not clear where a block of common questions should be placed in a question paper, given that it is advisable to place more difficult questions towards the end of the paper, and the common items should be either the easiest or the hardest questions depending on tier. It is possible that performance on common questions located towards the end of a foundation tier paper would be lower than if they were located towards the beginning of the paper because candidates had less time to answer them. Kolen and Brennan (2004) and Cook and Petersen (1987) note that common items should not appear in considerably different position on two tests else it might lead to items functioning differently in the two tests.

Rather than using a separate block of questions, M2 (Physics) and L2 (Additional Applied Science) presented common questions in approximately the same position in both tiers, avoiding ordering effects. This might have introduced context effects for the Physics paper, since the common items were often placed towards the end of multi-part questions in the foundation tier, and the beginning of multi-part questions in the higher tier.

Only one of the question papers surveyed (ICT) used several different question types, such as objective questions, short answer and extended answer questions. The ICT paper used constrained objective style questions on the foundation tier paper (with one exception) for non-common items, and used narrative short answer questions worth between two and six marks for the common items and higher tier specific questions. The remaining papers showed no differences in the choice of question type across the two tiers. Although it was tempting to conclude that differentiating the tiers by question type, as exemplified by the ICT paper, has contributed to the target-like patterning of common items across the tiers of this paper, it was difficult to draw firm conclusions on the basis of one paper. Further analysis of a wider range of question papers would be necessary to establish any trends.

We also investigated how similar items, which tested the same or similar content, varied across tiers. The similar items in L2 provided examples of ways in which similar items can be made easier for the foundation tier, despite testing similar content. Both foundation and higher tier items asked candidates to label a picture of a microscope. The wording and layout of the items were identical, except that foundation tier candidates were provided with a list of words from which to choose to label the microscope. This possibly was the reason why the question was much easier at the foundation tier (according to the Rasch estimates).

In the Mathematics paper H1 there was one pair of similar items, in which candidates were shown three scatter graphs and asked to describe the correlation shown in each diagram. The items were differentiated for the tiers by altering the scatter graphs, such that the different types of correlation were stronger for the foundation tier. Although this item was worth three marks on both papers, both foundation and higher tier candidates were asked to describe the correlation shown in each scatter graph. However, to receive full marks for this item, the higher tier had additionally to describe each correlation as strong or weak. Although it would have been possible to use the same layout for the item across tiers, there were differences between the tiers. The higher tier item provided space for candidates to respond immediately below each scatter graph. In contrast, for the foundation tier item each scatter graph was labelled, for example, Diagram 1, and candidates were asked to write their responses further below the scatter graphs, and link their response to the label given to each graph, rather than directly to the graph itself. It seems plausible that adding an additional step of linking responses to a label of a diagram rather than to the diagram itself would require more processing resources, because the label and the link to the actual diagram would need to be retained in working memory. The Rasch estimates for these questions demonstrated that the higher tier question was indeed more challenging, indicating that despite the difference in format between the foundation and higher tier, the difference in content made the higher tier item more difficult.

The qualitative review of items analysed the style, format and content of items in both the foundation and higher tier, with a particular focus on common and similar items. The analysis of individual items suggested that both question style and content play a role in the appropriate targeting of questions. Overall objective style questions seemed to be less challenging, as expected. The distinction between short answer and extended answer questions was less clear, although this may be due to the choice of question style targeted at each tier. For example, the ICT paper (H2) varied the style of questions between the tiers, from objective questions which featured only in the foundation tier, to common short answer questions, and extended answer questions in the higher tier only. However, a more extensive study of more question papers is necessary to determine whether this way of targeting questions to the higher and foundation tiers is effective. Examining questions which were similar, but not identical across tiers aimed to investigate how question structure and layout might contribute to the targeting of questions. However, it was striking that question structure and layout did not always relate to the degree of challenge posed by individual items. This is possibly because the questions investigated were well written and accessible for both tiers, so that the effect of modulation of question style across tiers was minimal. Instead, manipulating the content of similar questions across tiers seems to be of greater importance. This being the case, if questions assess the same content across tiers, it would be advisable to make such questions identical (common) across tiers to allow more effective evaluation of standards between tiers.

Discussion

We found that the grade C boundaries at the higher tier were, on average, close to the target set by OCR (40% of raw marks) whereas those at the foundation tier were, on average, considerably lower (64%) than the target (85% of raw marks). The lower-than-targeted grade C boundaries at the foundation tier explains why few component-pairs in this study were found near the target difference between the C boundaries (45 percentage points). To maintain standards across the tiers, the grade setting procedure should take into consideration the performance on common items. Currently, the emphasis is on maintaining year-on-year standards and the relative performance across tiers might not be given much weight. Where identical common items exist, and can be shown to have the same relative difficulty on each tier, vertical equating outcomes should be taken into consideration when setting common grade boundaries.

The selection of components in this study was based on the assumption that if the C boundaries on the foundation and higher tier were at a similar proportion of the paper total mark, there could potentially be an issue with the test construction (item writing). However, unexpected C boundaries such as this might also be obtained if the grade boundaries were not set appropriately during awarding. Dhawan (2012) presents a number of scenarios where the interaction

between the two issues – test construction and awarding – could lead to unexpected grade C boundaries. For instance, if the foundation tier was comparatively too difficult, it might lead to setting the C boundary very low to compensate. In the current study we focussed on test construction because if the items did not function as intended and an examination was harder or easier than it should have been according to the cohort ability, it would be appropriate to set lower or higher boundaries respectively to compensate. The focus, therefore, was on the review of item writing. In addition, the use of the other overlapping boundary in the two tiers, grade D, might have given slightly different results.

The comparison of the difficulty of the common items might be affected by context effects such as if the items were not in the same order in the two question papers. Ideally, we would want the common items to have the same stimulus and wording, position within a multi-part question, maximum marks and answer space. The mark schemes should also have the same wording and allow the same possible answers in each case. Along with the above criteria, we would expect that the common items were among the most challenging items in the foundation tier and the least challenging in the higher tier. If the tests were not designed keeping in mind these criteria, the consistent functioning of common items across tiers is likely to be adversely affected.

There are some caveats of equating tiered components. The results could be limited by the fact that equating is more appropriate for tests where the cohorts are not too different, whereas tiered examinations are targeted at cohorts expected to be different in ability. The strict set of assumptions for equating results to be adequate recommended by Kolen and Brennan (2004) is unlikely to be fully met in tiered examinations. However, the use of common items for equating is likely to provide more of a robust solution than some of the alternatives.

Large (positive) differences between the grade C boundaries of the two tiers might not be a foolproof indicator that the examinations were functioning as intended. However, comparing C boundaries is a simple procedure which can be carried out in each session. It can be used as an indicator of functioning of tiered components which can be explored further by a qualitative review of the questions. While it is easier to identify items which might not be functioning as intended using statistical evidence, pinpointing the actual cause of the inconsistent functioning could be challenging. Test development is a complex process – one which is influenced by many entities such as the curriculum, the item writers, the awarding bodies and the examinations held in the previous years. Although there are different sources of evidence available, the item writers are still required to 'predict' the difficulty of the items and target them at different grades. Writing of common items is even more challenging because it is expected that the same items should be appropriate in structure and format for both the foundation and higher tiers.

It is worth noting that, if the candidates were not correctly entered in the first place, a comparison of the tiered components is likely to be adversely affected. Future research in this area could focus on the actual process of how the candidates are entered in the tiers, who is involved, which factors are taken into consideration in making this decision, and how the entry decisions vary by different social indicators such as geographical region, gender and school type.

We explored some of the factors that could influence the relative functioning of the tiered examinations using statistical analysis and our perception of why some of the items might not be behaving as intended. Qualitative review of more components, possibly involving some of the item writers and subject experts, might give a better understanding of the functioning of the items. We found that the interpretation of the results was a demanding task because of the paucity of prior literature and specific guidelines – a challenge which the item writers might have to face as well. To conclude, we recommend that:

- a simple procedure such as comparing grade C boundaries could be carried out in each session to identify tiered components which might not be working as intended;
- the functioning of items could be investigated to check if the common items were indeed more difficult than the non-common items in the foundation tier and easier in the higher tier;
- where identical common items exist, and can be shown to have the same relative difficulty on each tier, vertical equating outcomes could be taken into consideration when setting common grade boundaries;
- the statistical evidence can feed into a qualitative analysis of questions to investigate if there were any concerns related to item writing;
- item writers should be provided with a set of specific and written guidelines for writing items in general and tiered examinations in particular;
- Ofqual could publish formal motivation for the decision to use differentiated assessment or not in a subject.

Acknowledgements

We would like to thank the Chairs of Examiners, Qualification Managers and Qualification Leaders of various subjects at OCR for their help and our colleague Tom Bramley for his advice.

References

Bond, T.G. & Fox, C.M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. 2nd Edition. New Jersey: Lawrence Erlbaum Associates, Inc.

Cook, L.L. & Petersen, N.S. (1987). Problems Related to the Use of Conventional and Item Response Theory Equating Methods in Less than Optimal Circumstances. *Applied Psychological Measurement*, **11**, 3, 225–244.

Dhawan, V. (2012). Monitoring the difficulty of tiered GCSE components using threshold marks for grade C. *Research Matters: A Cambridge Assessment Publication*, **14**, 18–21.

Good, F.J. & Cresswell, M.J. (1988a). Grade awarding judgements in differentiated examinations. *British Educational Research Journal*, **14**, 3, 263–80. Available online at: http://www.tandfonline.com/doi/pdf/10.1080/0141192880140304. (Accessed 22 February 2012).

Good, F.J. & Cresswell, M.J. (1988b). *Grading the GCSE*. Secondary Examinations Council: London.

Good, F.J. & Cresswell, M.J. (1988c). Placing candidates who take differentiated papers on a common grade scale. *Educational Research*, **30**, 3, 177–189. Available online at: http://www.tandfonline.com/doi/pdf/10.1080/0013188880300302. (Accessed 22 February 2012).

Hanson, B. A. & Béguin, A. A. (2002). Obtaining a Common Scale for Item Response Theory Item Parameters Using Separate Versus Concurrent Estimation in the Common-Item Equating Design. *Applied Psychological Measurement*, **26**, 1, 3–24.

Klein, L.W. & Kolen, M.J. (1985, April). *Effect of number of common items in common-item equating with non-random groups*. Paper presented at the annual meeting of the American Educational Research Association (AERA), Chicago.

Kolen, M.J. & Brennan, R.L. (2004). *Test equating, scaling, and linking. Methods and practices*. 2nd edition. New York: Springer-Verlag.

Morrison, C. A. & Fitzpatrick, S. J. (1992). Direct and indirect equating: A comparison of four methods using the Rasch model. *Research Bulletin 91–3*. Measurement and Evaluation Center, The University of Texas at Austin.

Ofqual (2011a). GCSE, GCE, Principal Learning and Project Code of Practice. Available online at: http://www.ofqual.gov.uk/for-awarding-organisations/96-articles/247-code-of-practice-2011. (Accessed 28 February 2012).

Ofqual (2011b).GCSE Subject Criteria for Classical Subjects. Available online: http://www.ofqual.gov.uk/downloads/category/192-gcse-subject-criteria. (Accessed 18 April 2012).

Ofqual (2011c). Review of Standards in GCSE English 2005 and 2009. Available online at: http://www.ofqual.gov.uk/files/11-09-22-Review-of-Standards-in-GCSE-English.pdf. (Accessed 18 April 2012).

Ofqual (2012a).GCSE Subject Criteria for History. Available online at: http://www.ofqual.gov.uk/downloads/category/192-gcse-subject-criteria. (Accessed 18 April 2012).

Ofqual (2012b). Standards over time. Available online at: http://www.ofqual.gov.uk/standards/92-articles/24-standards-over-time (Accessed 18 April 2012).

Ofqual (2012c). The new GCSE Examinations. Findings from the Monitoring of New Qualifications in French, Business and Geography 2010–11. Available online at: http://www.ofqual.gov.uk/files/2012-03-16-the-new-gcse-examinations.pdf?Itemid=144. (Accessed 18 April 2012).

QCA (2004). The statutory regulation of external qualifications in England, Wales and Northern Ireland. Available online at: http://www.ofqual.gov.uk/files/6944_regulatory_criteria_04(1).pdf. (Accessed 18 April 2012).

QCA (2005a). Review of standards in economics and business studies GCSE and A level 1998 and 2003. Available online at: http://www.ofqual.gov.uk/files/12889_econbusreport.pdf. (Accessed 18 April 2012).

QCA (2005b). Review of standards in biology GCSE 1998 and 2003; A level 1999 and 2003. Available online at: http://www.ofqual.gov.uk/files/12891_biologyreport.pdf. (Accessed 18 April 2012).

QCA (2005c). Review of standards in chemistry GCSE 1998 and 2003; A level 1999 and 2003. Available online at: http://www.ofqual.gov.uk/files/12890_chemistryreport.pdf. (Accessed 18 April 2012).

Wheadon, C. & Béguin, A. (2010). Fears for tiers: are candidates being appropriately rewarded for their performance in tiered examinations? *Assessment in Education: Principles, Policy & Practice*, **17**, 3, 287–300. Available online at: http://www.tandfonline.com/doi/pdf/10.1080/0969594X.2010.496239. (Accessed 22 February 2012).

Wright, B. D., & Stone, M. (1979). *Best Test Design*. Chicago: MESA Press.

Statistical Reports

The Research Division

The on-going 'Statistics Reports Series' provides statistical summaries of various aspects of the English examination system such as trends in pupil uptake and attainment, qualifications choice, subject combinations and subject provision at school. These reports, produced using national-level examination data, are available on the Cambridge Assessment website: http://www.cambridgeassessment.org.uk/ca/Our_Services/Research/Statistical_Reports.

The most recent additions to this series are:

- Statistics Report Series No.47:
  Uptake of two-subject combinations of the most popular A levels in 2011, by candidate and school characteristics.
- Statistics Report Series No.48:
  A Level Uptake and Results, by Gender 2002–2011.
- Statistics Report Series No.49:
  GCSE Uptake and Results, by Gender 2002–2011.
- Statistics Report Series No.50:
  A Level Uptake and Results, by School Type 2002–2011.
- Statistics Report Series No.51:
  GCSE Uptake and Results, by School Type 2002–2011.

Research News

Conferences and seminars

Society for Research into Higher Education (SRHE)

In December 2012, Irenka Suto attended the SHRE conference in Newport, Wales. The conference focused on a key question currently being asked about higher education across the globe: What is Higher Education for? Irenka presented on *'How well prepared are new undergraduates for university study? An investigation of lecturers' perceptions and experiences'*.

Seventh Annual UK Rasch Users meeting

Tom Bramley attended the seventh UK Rasch Users meeting at the University of Manchester in March. The Group provides a forum for Rasch enthusiasts working in different fields to get together to share ideas and present research. Tom presented a paper entitled *'A Rasch-based method for back-filling a specification grid for an exam paper'*.

Publications

The following articles have been published since Issue 15 of *Research Matters:*

Benton, T. (2013). Exploring equivalent forms reliability using a key stage 2 reading test. *Research Papers in Education*. Available online at: http://www.tandfonline.com/doi/abs/10.1080/02671522.2012.754227

Bramley, T. & Dhawan, V. (2013). Problems in estimating composite reliability of 'unitised' assessments. *Research Papers in Education*. Available online at: http://www.tandfonline.com/doi/abs/10.1080/02671522.2012.754226

Crisp, V. (2013). Criteria, comparison and past experiences: How do teachers make judgements when marking coursework? *Assessment in Education: Principles, Policies and Practice*. Available online at: http://www.tandfonline.com/doi/abs/10.1080/0969594X.2012.74105920

Dhawan, V. & Bramley, T. (2012). Estimation of inter-rater reliability. *Ofqual Reliability Programme Report*. Office of Qualifications and Examinations Regulation. Available online at: http://www.ofqual.gov.uk/files/2013-01-17-ca-estimation-of-inter-rater-reliability-report.pdf

Johnson, M. & Shaw, S. (2012). Interpreting examiners' annotations on examination scripts: a sociocultural analysis. *Irish Educational Studies*. Available online at: http://www.tandfonline.com/doi/abs/10.1080/03323315.2012.739828

Johnson, M. & Lewis, C. (2012). 'Can you dig it?': developing an approach to validly assessing diverse skills in an archaeological context. *Journal of Vocational Education and Training*. Available online at: http://www.tandfonline.com/doi/full/10.1080/13636820.2012.755212

Johnson, M. & Black, B. (2012). Feedback as scaffolding: Senior Examiner monitoring processes and their effects on examiner marking. *Research in Post-Compulsory Education*. Available online at: http://www.tandfonline.com/doi/abs/10.1080/13596748.2012.738965

Shaw, S. (2012). International Assessment of Geography through the medium of English: analyzing the language skills required. *Bilingual teaching – globalization, regional Geography and English integration*. **16**, 24–44.

Shaw, S., Johnson, M. & Warwick, P. (2012). The Assessment for Learning in International Contexts (ALIC) Research Project: Understanding Assessment for Learning: values and practices across diverse contexts. *Research Intelligence*, **119**, 14–16.

Vidal Rodeiro, C.L. & Nadas, R. (2012). Effects of modularity, certification session and re-sits on examination performance. *Assessment in Education: Principles, Policies and Practice*. Available online at: http://www.tandfonline.com/doi/abs/10.1080/0969594X.2011.614218

Back Issues

Research Matters is a free biannual publication from Cambridge Assessment. The aim of this publication is to share assessment research in a range of fields with the wider assessment community and to comment on prominent research articles.

The first issue of *Research Matters* was published in September 2005 and we have published 16 issues to date. Special Issues reporting on particular research that requires a longer and more detailed publication have also been published. The first of these, *Variations in Aspects of Writing in 16+ English Examinations between 1980 and 2004*, was published in November 2005. A second *Research Matters* Special Issue, *Comparability*, was published in October 2011, and in October 2012 we published a third Special Issue, *An approach to validation*.

Full details of the contents, articles and features of all previous issues of *Research Matters* are available on our website: http://www.cambridgeassessment.org.uk/ca/Our_Services/Research

A limited number of back issues of *Research Matters* are available on request. Please email: researchprogrammes@cambridgeassessment.org.uk stating the issue number and/or copy of the Special Issue you are interested in. Due to availability, requests will be limited to three back issues per person/organisation. All previous issues are available to download in full from our website.

Requests from readers wishing to be added to our mailing list to receive regular copies of all future issues of *Research Matters* are also welcome at the above email address.

CONTENTS : Issue 16 June 2013

Cambridge Assessment
1 Hills Road
Cambridge
CB1 2EU

Tel: 01223 552666
Fax: 01223 552700
Email: ResearchProgrammes@cambridgeassessment.org.uk
http://www.cambridgeassessment.org.uk

ISSN: 1755–6031