

The evaluation of the impact of changes to A levels and GCSEs

Second interim report

Alpha*Plus* Consultancy Ltd

This research report was commissioned before the new UK Government took office on 11 May 2010. As a result the content may not reflect current Government policy and may make reference to the Department for Children, Schools and Families (DCSF) which has now been replaced by the Department for Education (DFE).

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

Executive summary	4
Introduction	4
Aims of the evaluation	4
Background to the evaluation	5
Methodology	5
Data sources	5
Summary of preliminary findings	6
The impact of changes to the specifications on teaching and learning	6
The impact of the assessment regime on student development	7
The impact of centre approaches on students	9
Attainment in the context of the new specifications.....	10
Determining factors for driving educational change	11
1 Introduction	12
1.1 Background to the evaluation	13
1.2 Aims of the evaluation	14
1.3 Data sources	14
1.4 Scope and limitations of this report	15
2 Methodology	15
2.1 Centre-based case study data	15
2.2 Wider stakeholder data	15
2.2.1 Interviews with stakeholders	15
2.2.2 Higher education institution survey	16
2.3 National datasets used	17
2.4 Centre online survey	18
3 A level findings	19
3.1 The impact of changes on teaching and learning	19
3.1.1 Maintaining standards over time	20
3.1.2 The impact of specification changes on student attainment at A2	22
3.1.3 Introducing stretch and challenge	25
3.1.4 The development of subject-specific skills and knowledge.....	28
3.1.5 Opportunities for higher-level generic skills development.....	29
3.1.6 Participation and student engagement	31
3.1.7 Changes in participation	33
3.1.8 Preparation for progression	34
3.2 The impact of change on centre behaviour	36
3.2.1 Choice of specification	36
3.2.2 Management and resource implications for centres	37
4 GCSE findings	41
4.1 The impact of changes on teaching and learning	41
4.1.1 The development of subject-specific skills and knowledge.....	41
4.1.2 Opportunities for higher-level generic skills development.....	42
4.1.3 Participation and student engagement	43
4.1.4 Preparation for progression	44
4.1.5 Attainment.....	44
4.1.6 Impact of controlled assessment on teaching and learning	45
4.2 The impact of change on centre behaviour	48
4.2.1 Influence on centre choice of specification	48

4.2.2	Management and resource implications for centres	49
5	Wider stakeholder perceptions of the impact of specification changes..	53
6	Summary of preliminary findings	56
6.1	The impact of changes to the specifications on teaching and learning.....	56
6.2	The impact of the assessment regime on student development	57
6.3	The impact of centre approaches on students	60
6.4	Attainment in the context of the new specifications	60
6.5	Determining factors for driving educational change.....	61

Executive summary

Introduction

This report records findings from further data collection and analysis for the independent evaluation on the impact of changes to A levels and GCSEs in England,¹ which took place in autumn 2010 and spring/summer 2011. In the course of this evaluation, initial perceptions were revisited, and data on attainment for the first A level results for the new specification assessments was examined. The report builds on the baseline of statistical data developed in June 2010 and on the findings from fieldwork carried out before the publication of results from the summer 2010 examinations.² The earlier findings were originally reported on in a First Interim Report (July 2010).³ The executive summary for the First Interim Report submitted is included as Appendix 1 to the main report.

A final report, due in January 2012, will chart the process of change and impact that has been seen in centres over the course of the evaluation. This second interim report provides a summary of preliminary findings in advance of the final data collection, analysis and reporting.

Aims of the evaluation

The evaluation focuses on the effect of the changes to A levels and GCSEs on teaching and learning and on the management of assessment. This report focuses on building an understanding of the impact of the changes to specifications for A levels and GCSEs at this stage of implementation – in other words, on the basis that the specifications have been implemented at incremental points within a two-year timeframe and that interpretation of the data may therefore depend on where stakeholders are located on the change continuum for each qualification and/or subject.

¹ Changes in all A level subjects, except mathematics, were introduced in September 2008. These changes included a move from six units to four for the majority of subjects, the introduction of greater stretch and challenge at A2, and the A* grade. Changes to GCSEs in all subjects except English, mathematics, ICT (information, communication and technology) and the sciences were introduced from September 2009; changes to GCSE English, mathematics and ICT were introduced in September 2010. At GCSE, controlled assessment and unitisation were widely introduced.

² The first round of data for the evaluation was collected at the end of a full cycle of A2, and two cycles of AS, before the summer 2010 examinations; data for GCSE was collected towards the end of the first year of a two-year course.

³ The independent evaluation of the impact of the changes to A levels and GCSEs was commissioned by QCDA in March 2010. The contract was transferred to the Department for Education (DfE) in January 2011. The evaluation was subject to a prolonged state of 'purdah' before and after the General Election in May 2010, during which time communication with stakeholders was limited to work that had already commenced.

Background to the evaluation

The initial round of data collection during 2010 concentrated on six GCSE subjects (French, geography, health and social care, history, media studies and Spanish – for which new specifications were introduced in September 2009) and on six A level subjects (English literature, French, geography, media studies, physics and psychology). A change of focus⁴ following the Department for Education's (DfE) take-over of the contract for the evaluation has meant that, at GCSE, data was collected for English/English language and mathematics (for which new specifications were introduced from September 2010) instead of for health and social care and media studies. At A level, data was collected for history instead of media studies.

Methodology

Data sources

This report uses the following data sources:

- Awarding Body Data Archive (ABDA) – awarding organisation data supplied by Ofqual
- Official Statistics – Statistical First Releases (SFR) – from the DfE Research and Statistics Gateway
- National Pupil Database (NPD) – including Census, Pupil Level Annual School Census (PLASC) and examination results
- case study data from 17 case-study centres (across two rounds of case study visits)
- wider stakeholder interview data – 18 interviews with wider stakeholder representative groups, including 2 awarding organisations, and 35 responses to an online survey from 29 English higher education institutions (HEIs)⁵

⁴ The new GCSEs added to the evaluation (English and mathematics) are included in the national curriculum for Key Stage 4 (2007). The change of focus at GCSE and A level was made at the request of the DfE, to reflect the coalition government's interest in the 'core' subjects.

⁵ An invitation to take part in an online survey was sent to 103 HEIs in England and followed up with reminder emails and telephone calls; Supporting Professionalism in Admissions (SPA) also promoted the survey at one of their events and in a newsletter. Response rate of 27% from the original 103 HEIs invited – one additional HEI also responded.

-
- awarding organisation and Ofqual document review (updated in 2011 to include new subjects)⁶
 - centre online survey data⁷ – responses from a total of 44 centres (6.4% response rate): 19 completed responses to a senior leader questionnaire, 97 responses from 41 different centres to the subject-specific questionnaires.

Summary of preliminary findings

The impact of changes to the specifications on teaching and learning

Just over half of the wider stakeholders interviewed expressed concern that centres had not recognised the degree of change in the A level specifications – and some had particular concerns about the greater stretch and challenge required at A2 – and the subsequent consequences for teaching and learning.

‘Stretch and challenge’ is being interpreted in a variety of ways, and there is a degree of uncertainty and ambiguity about what is involved. Overall, however, the majority of centres felt that there was greater stretch and challenge in the new specifications at A2, although this was not always equally reflected in the assessments. Physics was the only subject for which centres felt that there was less stretch and challenge than in the previous specifications.

It was reported that, to develop higher-level skills and independent learning in students, teachers need time to adjust their teaching strategies and students need time to develop new approaches in their learning.

The decrease in the number of units at A level from six to four (a decrease that does not apply to physics) was seen as positive in terms of ensuring greater depth and breadth of study.

Progression from AS to A2 was considered by many case-study centres to be a larger jump than in the previous specifications. During the follow up visits in 2011, for example, for English literature, history and geography those who responded thought that the AS was preparing students less well for A2 than before

⁶ The review of the awarding organisation documentation follows the logical development process of the revised qualification: the original qualification criteria and the resulting awarding organisation specification(s); the revised subject criteria for each qualification and the responding awarding organisation specification(s); the sample assessment materials, along with relevant mark schemes, and any other qualification-specific awarding organisation guidance material aimed at teachers and learners.

⁷ The general qualifications centre survey was delivered as nine separate online questionnaires. The survey was available from 10 May 2011 to 10 June 2011. The nine questionnaires were: four combined A level and GCSE questionnaires – English, French, geography, history; two A level-only questionnaires – psychology, physics; two GCSE-only questionnaires – mathematics, Spanish, and one generic senior leaders’ questionnaire.

Teachers felt that A2 English literature, geography, psychology, history and French were more challenging, and considered them better preparation for higher education – for example, in terms of encouraging independent learning. There were some exceptions in terms of specific skills development for a subject – for example, some geography teachers felt that there was not enough emphasis on the research skills required for university study.

Although the updating of course content had increased students' sense of motivation and engagement with GCSE subjects, this positive effect has been counteracted by the increased focus in teaching and learning on assessment.

The impact of the assessment regime on student development

A number of centres commented on the rise in numbers of A level students re-sitting examinations in order to achieve a higher grade. Students generally welcomed the opportunity for re-sits, with the majority believing re-sits should be allowed in any circumstance.

Although coursework was considered a burden for teachers in terms of their workload, it was also felt by many teachers to offer assessment which supported development (i.e. with a formative and summative element). Teachers, managers and students suggested that A level coursework performance was a better indicator of a student's ability than written examinations. Physics apart, the majority of A level students' subject- and non-subject-specific comments suggested that they would wish to retain coursework as part of the assessment regime.

There were some concerns about unitised assessment for GCSE. Five case-study centres discussed the respective merits of unitised and linear assessment in GCSE French, history and mathematics. They were of the view that linear assessment promoted more in-depth and longer-lasting learning than unitised assessment, particularly in relation to the development of subject-specific skills. Similar concerns were expressed by wider stakeholders. Unitised assessment with opportunities for retakes and early entry, coupled with pressure on centres and students to get results 'under their belt', was considered at odds with the value placed on synoptic assessment and on the development of skills rather than a narrow knowledge-based approach to subjects. However, this was balanced to some extent by teachers and students who felt there were benefits in allowing students to achieve along the way and to build on results.

The overall consensus from centres and awarding organisations was that there is conflict between the need to gain the 'results' required for school performance targets – using re-sits and early entry to maximise pass rates (e.g. at A*–C at GCSE), which is possible with

unitised qualifications – and encouraging learning, development, coherence and a greater understanding of the subject.

Some case-study teachers reported that students are very outcome-focused and often reluctant to engage with any content that they do not think will be in the assessments. Stakeholders felt that the focus on assessment and assessment strategy diverted time and attention from high-quality teaching.

A wide range of approaches to controlled assessment⁸ is being implemented within and across centres, often depending on the subject being assessed and/or on the teacher. Awarding organisations have often interpreted Ofqual's generic guidance differently, meaning that there is sometimes variation in the advice provided for the same subject across the different awarding organisations. This has resulted in considerable variation in the amount of support that teachers believe they can give their students in the preparation stage of controlled assessment.

Concern was expressed that some teachers are using strategies, even if often unwillingly, to help students to get the best grade, and that these undermine the validity of the controlled assessments in terms of the knowledge and skills that are meant to be assessed. Examples of strategies include running practice assessments that change little in the 'real' assessments and modern foreign language teachers setting more controlled assessments than required and choosing the best for each individual student.

About half of the centres that had previously found that the scheduling of controlled assessments involved a significant added management burden reported now that they were coping well or satisfactorily with the issue. The remaining half thought that the challenge of fitting controlled assessments into the school calendar was becoming a more serious issue for them, as the number of departments involved was increasing. The majority of controlled assessment is undertaken during subject lesson time in many centres, although for modern foreign languages the need for assessment on a one-to-one basis means that students are sometimes taken out of other subject lessons. Many of the centres reported problems when there is a need to rearrange a controlled assessment for students who have been absent from the scheduled assessment, or when a student's behaviour means that they have to be asked to leave the room and miss part of the assessment.

It is accepted that change, such as the introduction of controlled assessment, initially increases teachers' workload and that this effect is usually only short term. However, teachers of modern foreign languages reported that they expect their workloads to remain

⁸ Controlled assessment is a new form of internal assessment of the work of a course, replacing coursework.

heavier in the longer term also, because of specific requirements to do with organising and conducting controlled assessment.

Most teachers liked the opportunity that controlled assessment gives to choose topics and contextualise the tasks for their students. However, there was no consensus across or within centres and subject areas about the extent to which controlled assessment is an appropriate form of assessment for particular skills and knowledge. Controlled assessment was seen as summative assessment, with only a minority of teachers seeing it as having a formative role and being an integral part of the teaching and learning process. There are concerns that, rather than promoting in-depth independent learning, controlled assessment may only test the ability to learn content and regurgitate it.

When students talked about controlled assessment in terms of a specific subject, they tended to be negative – but when they talked generally about it, they were very positive. Students of history and geography were the most positive about controlled assessment, and students of Spanish and French the most negative. History and geography students liked the fact they could use their research notes during the controlled assessment. The majority of students also stated that they liked the fact that controlled assessment did not take place in the main school hall, which is where they sat external examinations, and that controlled assessment was therefore less stressful. There was some belief that controlled assessment can bring advantages when properly implemented and that a measured approach should be taken to revisions of the system. Staff at all levels expressed the need for guidance on best practice in the management of controlled assessment. Although awarding organisations got some credit for providing information and advice, a few teachers regretted that new information was provided only when controlled assessment was already underway, which led to some confusion.

The impact of centre approaches on students

There is evidence from many of the case-study centres to suggest that the focus on school performance measures is a strong driver for centres to ensure students meet target grades. In many cases this means that teaching and learning is driven by the assessment regime, with centres relying heavily on awarding organisations for training to support students to pass examinations. There is evidence to suggest that some teachers and students will not consider specific content and/or skills that are not expected to be assessed as a priority, but rather as an indulgence. In some centres, however, the focus is less narrow, with centres embracing the change and challenge presented.

Choice of specification was reportedly determined by several factors, including the appropriateness for the needs and aspirations of students, consistency with the nature of the subject and familiarity with the awarding organisation.

Attainment in the context of the new specifications

The A* grade at A level is perceived as recognising high attainment by teachers and students. Some universities are starting to ask for an A* pass for certain courses.

Centres reported the use of a range of models for the management of the extended project qualification. The qualification was felt to have benefits for both student and staff development. Wider stakeholders expressed enthusiasm for the higher-level skills developed as part of the extended project qualification.

The plateau in grades at A level seen in the statistical data⁹ does not fully reflect the increased level of challenge perceived by students and teachers before results from the summer 2010 examinations, especially at A2, where many teachers had expected to see a drop in attainment after the change of specification. The plateau in grades is likely to be due to measures introduced by Ofqual to ensure that the first candidates to complete the new specification A levels did not suffer disadvantage as a result.

Analysis of demographics, prior attainment and the impact of re-sitting in the statistical data does not offer an explanation for the overall plateau in A level grades for the new specifications that occurred in summer 2010. Changes in the type of candidate taking the new specification A levels may provide the most obvious explanation for a change between that year and previous years in terms of grades attained. The plateauing of grades from 2009 to 2010, however, cannot be explained by factors relating to the candidates' background. Historically A level grades have been increasing year on year, with a similar trend of increase to the grades achieved by candidates at GCSE; and yet, if the GCSE results of 2010 A level candidates are compared with those taking A levels in 2009 and 2008, the 2010 candidates' GCSE results are similarly higher, but the A level grades they

⁹ From 1996 to 2009 (the last year in which the previous GCE specifications were awarded), A level attainment continued broadly on the same improving trend in most subjects – a small steady increase of about 0.06 of a grade per year in grades attained, punctuated by short-term dips downwards from the trend at times of specification change. Also, since around 2003, participation has been increasing in most subjects.

For results in 2010, a year in which another specification change has occurred (the specification change for teaching from 2008 working through to results in 2010), attainment saw a plateauing of grades following a long period of increasing grades – i.e. another shift downward from the trend, reminiscent of results in previous change periods in curricula (although much smaller in size than the dip in 2000/01). Looking at individual grades attained (across all new specification subjects combined), the same proportion of students obtained the A grade (including those who received the newly introduced A* grade) in 2010 as received an A grade in 2009 (24.6%), whereas, for example, from 2003 to 2009, on average year on year 0.7% more students received an A grade.

attained have not increased. In addition, the demographic make-up of the 2010 A level candidature is not demonstrably different from that of the 2009 candidature.

Concerns about an expected drop in attainment suggest that centres may not fully understand how the awarding principles of maintaining standards are intended to support a gradual change; they also suggest that awarding organisations are required to manage change in ways that do not disadvantage a particular cohort of students (i.e. those who are facing the changes for the first time).¹⁰

Determining factors for driving educational change

The impact and speed of change seen to date is determined in part by differences in:

- the stage of implementation – i.e. whether centres have already completed a full cycle of the new specification
- the extent to which centres recognised change in the specifications and in the method of enquiry these encourage
- where teachers and students were in terms of adjusting to new approaches to teaching, learning and development
- how centres and/or individual teachers perceived the role of assessment in teaching and learning and, in particular, the extent to which teachers felt the assessments were compatible with, and valid for, the subject and the skills to be assessed
- how pressured centres and teachers felt to ensure that, overall, their students gained the desired pass rates and grades to meet national targets
- expectations and perceptions of achievement for particular student cohorts and/or profiles
- how confident centres felt in managing the change process, their capacity, and the support offered and taken up.

It is possible that these differences would be the same for any curriculum and/or specification change and, overall, would be expected in these early stages of implementation. However, there is strong evidence to suggest that new specifications alone cannot drive educational reform. A range of other drivers and context related variables such as the focus on assessment and performance targets are impacting on centre behaviour in terms of teaching, learning and development, and ‘skew’ classroom practice.

¹⁰ The prime objectives of maintaining grade standards over time and across different specifications within a qualification type (Ofqual, April 2010, *GCSE, GCE, Principal Learning and Project Code of Practice*, para 6.2) necessarily become more problematic, and concerns among stakeholders more likely to arise, at times of curricular change.

Explanations for the plateau in A level grades awarded in 2010 can only be tentative at this stage. As teachers and students adjust to new approaches in teaching and learning, further statistical data analysis is needed, following the summer 2011 assessment results, to see if this plateau is a result of the 'newness' of the specifications. Additional data from centres will allow a more detailed exploration and understanding of the results in terms of specific cohort groups, including students who need longer to develop independent learning skills or greater support with conceptualised approaches.

Although there is some suggestion from centres that the new specification at AS is discouraging progression to the full A level in specific subjects at a local level, further analysis of the statistical data is needed to see the extent to which this is reflected nationally.

A more complete picture of attainment for GCSE will be possible following analysis of the summer 2011 statistical data. The high level of interest in the impact of unitisation and controlled assessment on teaching and learning has made this a key focus of this report. Fieldwork in the autumn term 2011 will offer an opportunity to look in more detail at the impact of other changes to the specifications.

1 Introduction

This report gives interim findings from the independent evaluation on the impact of changes to A levels and GCSEs in England. At A level, these changes were introduced in September 2008, to all A level subjects, except mathematics. The changes included a move from six units to four for the majority of subjects, the introduction of greater stretch and challenge at A2, and the A* grade. Changes to GCSEs were introduced from September 2009 in all subjects except English, mathematics, ICT (information, communication and technology) and the sciences; changes to GCSE English, mathematics and ICT were introduced in September 2010. At GCSE, controlled assessment and unitisation were widely introduced.

The report follows two rounds of fieldwork with case-study centres and wider stakeholder groups. It builds on the baseline of statistical data developed in June 2010 and reports on early indications of impact and issues based on the understanding, perceptions and attitudes of centres and wider stakeholder groups.

The first round of data for the evaluation was collected at the end of a full cycle of A2, and two cycles of AS, before the summer 2010 examinations; data for GCSE was collected towards the end of the first year of a two-year course. The data was originally reported on in

the First Interim Report (July 2010).¹¹ Data collected in 2010 focused on stakeholders' perceptions of the impact of the changes, which were based on an emerging picture. These findings offered evidence-based insights into the initial and short-term impact of changes to qualification specifications on centres, students, awarding organisations and wider stakeholder groups. The executive summary for the First Interim Report is included here as Appendix 1. The second round of data collection took place in autumn 2010 and spring/summer 2011, revisiting initial perceptions, and looked at statistical attainment data for the first A level results for the new specification assessments.

Final reporting will be in January 2012 after a full cycle of both qualifications and will synthesise findings to date, together with statistical attainment data from the summer 2011 examination series and a further round of interviews with case-study centres and stakeholders. The final report will chart the process of change and the impact seen in centres during the complete evaluation period (March 2010–December 2012).

This second, interim report provides a summary of preliminary findings ahead of this process of final data collection, analysis and reporting.

1.1 Background to the evaluation

The initial round of data collection during 2010 concentrated on six GCSE subjects (French, geography, health and social care, history, media studies and Spanish – for which new specifications were introduced in September 2009) and on six A level subjects (English literature, French, geography, media studies, physics and psychology). A change of focus¹² following the Department for Education's (DfE) take-over of the contract for the evaluation has meant that, at GCSE, data was collected for English/English language and mathematics (for which new specifications were introduced from September 2010) instead of for health and social care and media studies. At A level, data was collected for history instead of media studies.

¹¹ The independent evaluation of the impact of the changes to A levels and GCSEs was commissioned by QCDA in March 2010. The contract was transferred to the Department for Education (DfE) in January 2011. The evaluation was subject to a prolonged state of 'purdah' before and after the general election in May 2010, during which time communication/fieldwork with stakeholders was limited to work that had already commenced. This meant that an online survey of centres planned was not permitted.

¹² English and mathematics, the new GCSEs added to the evaluation, are included in the national curriculum for Key Stage 4 (2007). The change of focus at GCSE and A level was made at the request of the DfE, to reflect the coalition government's interest in the 'core' subjects.

1.2 Aims of the evaluation

The evaluation focuses on the effect of the changes to A levels and GCSEs on teaching and learning and on the management of assessment.

This effect is considered under three themes:

- stakeholders' perceptions of the new GCSEs and A levels
- the impact of the changes on teaching and learning
- the impact of the changes on centre behaviour and management of change (including assessment).

This report focuses on building an understanding of the impact of the changes to specifications for A levels and GCSEs at this stage of implementation – in other words, on the basis that the specifications have been implemented at incremental points within a two-year timeframe and that interpretation of the data may therefore depend on where stakeholders' are located on the change continuum for each qualification and/or subject.

Questions addressed under each theme can be found in Appendix 2.

1.3 Data sources

The interim report uses the following data sources:

- Awarding Body Data Archive (ABDA) – awarding organisation data supplied by Ofqual
- Official Statistics – Statistical First Releases (SFR) – from the DfE Research and Statistics Gateway
- National Pupil Database (NPD) – including Census, Pupil Level Annual School Census (PLASC) and examination results
- case study data from 17 case-study centres – including interviews or focus groups with senior leadership teams, heads of department, subject teachers, examination officers, and students
- interviews with 18 wider stakeholder representative groups, including 2 awarding organisations, and 35 responses to an online survey from 29 English higher education institutions (HEIs)
- awarding organisation and Ofqual document review (updated in 2011 to include new subjects)¹³

¹³ The review of the awarding organisation documentation follows the logical development process of the revised qualification: the original qualification criteria and the resulting awarding organisation specification(s); the revised subject criteria for each

-
- centre online survey data – 44 centres responded in total: 19 completed responses to the senior leader questionnaire, 97 responses from 41 different centres to the subject-specific questionnaires.

1.4 Scope and limitations of this report

This report includes data on full GCSEs only. The original round of data collection (commissioned by the Qualifications and Curriculum Development Agency – QCDA) included further education (FE) centres as wider stakeholders. For the second round of data collection (following transfer of the evaluation to the DfE), FE centres have been included as providers of full-time A level provision for 16–18 year olds. However, the scope of the evaluation does not currently include any additional statistical datasets specific to this cohort, as FE attainment data has not been identified as a priority for the resource available.

2 Methodology

2.1 Centre-based case study data

The original 15 case studies visited in May/ June 2010 were revisited in spring 2011, with the exception of two centres, which were replaced by an FE college and a secondary school (both of which had sixth forms). Details on how the case study sample was selected can be found in Appendix 3.1.1, together with a breakdown of centre characteristics for the second round of data collection. Overall, there were visits to 17 case-study sites.

The visits included interviews with senior leadership teams, heads of department, subject teachers, examination officers and student focus groups. An overview of subject and role coverage of the staff interviewed across the two visits is included in Appendix 3.1.2.

2.2 Wider stakeholder data

2.2.1 Interviews with stakeholders

The activity reported here straddles the two main rounds of data collection. An initial group of wider stakeholders was interviewed before July 2010. Further data was collected from an additional sample of wider stakeholders from August to October 2010 and from a follow-up meeting with one of the FE representative groups in June 2011 and an initial meeting with two awarding organisations.

Stakeholder telephone interviews or face-to-face group interviews were held with an agreed sample from the following groups: employers, subject and professional associations, FE representative organisations, trade unions, workforce development agencies and awarding organisations.

The interests of the wider stakeholder group were varied and the evaluation themes were therefore not equally relevant to each stakeholder. The research tools developed were based on topic areas and were used flexibly, depending on the focus of the wider stakeholder organisation interviewed.

The findings were drawn from the analysis of 18 wider stakeholder interviews. Because they either had limited availability or felt they knew too little for their responses to be of any merit, eight further wider stakeholder organisations that were approached felt unable to contribute, and five subject associations declined to be interviewed. One subject association felt unable to comment until it had received feedback from a membership survey; the same organisation requested payment for data in recognition of the charitable status of the organisation. Two workforce development agencies considered themselves unsuitable for inclusion in the evaluation, one stating that the topics fell outside the remit of what their organisation delivers on behalf of the DfE and referring to the uncertain political climate and to pressures of work. Employer organisations generally considered themselves insufficiently familiar with the changes to comment. One was unable to comment because of an internal dispute between training and marketing.

It is recognised that the sample represents a relatively narrow stakeholder perspective. Perceptions of the impact of the changes to A levels and GCSE will be influenced by the different agendas of different organisations.

2.2.2 Higher education institution survey

An invitation to take part in an online survey was sent to 103 HEIs in England and followed up with reminder emails and telephone calls; Supporting Professionalism in Admissions (SPA) also promoted the survey at one of their events and in a newsletter. The survey ran from 23 March 2011 to 27 May 2011. Overall, 35 responses were received from 29 HEIs (seven from different departments at one university).¹⁴ Of these, 20 of the respondents were from admissions or student recruitment, 10 from individual departments or colleges, and a further 5 from registry; 26 of the respondents had a job role which specifically mentioned admissions or recruitment; 27 of the respondents described their role as predominantly managerial or administrative, 8 as predominantly academic.

¹⁴ Response rate of 27% from the original 103 HEIs invited – one additional HEI also responded.

Of the 29 institutions responding:

- there were 15 pre-1992 universities and 11 post-1992 universities, with two university colleges and one private provider; 52% of the HEIs in the survey are therefore pre-1992 universities, against a national figure of 49%
- 7 of the 29 HEIs were Russell Group Universities, which means that Russell Group Universities are slightly over-represented in the survey (24% of survey HEIs against 15% nationally – there being 16 English members of the Russell Group)

This means that findings should not be generalised to the whole population of HEIs.

2.3 National datasets used

The work undertaken in summer 2010 aimed to establish a baseline dataset for results that pre-dated the first results from the new-specification A level results. Analysis for this report looks at the impact of the new specifications on participation, progression and attainment on the basis of the first full suite of results. As such, the investigations are focused on change (expected and unexpected within those three main measures). As for the previous cycle of analysis, the investigation is based on the following three datasets.

ABDA data provided by Ofqual (previously provided by QCDA) – this includes GCSE and A level exam results from Key Stage 4 (KS4) and Key Stage 5 (KS5). This data, which was anonymised, included unit performance, prior attainment and re-sit information. Starting in 2008, with GCSE and A level, results have been collected in alternate years, for only a sample of subjects, and for only a single specification/specification strand for each subject (usually the specification with the highest candidate volume, where awarding organisations provide more than one specification for a subject). This subject selection aligns closely, however, with the sample to be investigated in the evaluation. Across the subjects considered in this report, the ABDA data accounts for around 60% of all candidates taking the subjects. This data is identified as ‘ABDA’ in the report.

Official Statistics: DfE Research Gateway Statistical First Release Data¹⁵ – these datasets provide aggregate entry and achievement data for individual subjects (consolidated for all awarding organisations) and for the key school targets (e.g. five A*–C grades incl. English & mathematics) for candidates completing KS4 and KS5. They are essentially pre-run reports on the NPD dataset but are considered separately here, because the quality

¹⁵ From the DfE Research Gateway Statistical First Release Data: <http://www.education.gov.uk/rsgateway/index.shtml>

assurance of those reports is undertaken by the DfE rather than the evaluation team. This data is identified as 'Official Statistics' or Official Statistics '(SFR)' in the report.

DfE National Pupil Database (NPD) – the DfE provides the evaluation team with a subset of these large datasets, including KS3, KS4 and KS5 results and a range of demographic indicators about students and the schools they attend. The datasets provide identifiable candidate data, thereby allowing matching across the datasets for progression analysis. This data is referred to as 'NPD' in the report.

2.4 Centre online survey

The centre survey was delivered as nine separate online questionnaires. The survey was available from 10 May 2011 to 10 June 2011. The nine questionnaires were:

- four combined A level and GCSE questionnaires – English, French, geography, history
- two A level-only questionnaires – psychology, physics
- two GCSE-only questionnaires – mathematics, Spanish
- one generic senior leaders' questionnaire.

Most of the questions posed in the eight subject-specific questionnaires were essentially the same for each subject/level but, where appropriate, the questions were specifically worded for the subject. The questionnaires for A level only and for GCSE only were essentially subsets of the combined A level and GCSE questionnaires. The questions generated quantitative (nominal, ordinal and ratio) and qualitative data.

A total of 691 centres were invited to take part in the survey, of which 44 responded to one or more questionnaire.¹⁶ For the senior leader questionnaire, there were 19 completed responses to the questionnaires, each from a different centre. For the subject-specific questionnaires, there were 97 responses to the questionnaires from 41 different centres. There was a relatively low response rate (6.4%) to this survey despite reminder telephone calls, letter and emails to centres. Centres that did not respond cited the timing (pre-summer examinations) and existing heavy workloads as reasons for not doing so. An overview of centre characteristics for participating centres is included in Appendix 3.2. Although the low

¹⁶ An initial stratified sample was identified from Edubase (whole population listed as 4,007 centres, although some of these were later removed after further research had identified them as inappropriate – for example, because they were consortia, or adult education centres or because the centre had closed). Further centres were identified using a purposive strategy if there was a gap in the data available via Edubase e.g. FE colleges. Strata used were transition (to age 16; through 16; 16+), prior attainment (selective; non-selective), level of deprivation (calculated on the basis of the percentage of students entitled to free school meals), teaching and learning (in special measures; not in special measures).

level of response to the online survey means that any findings should be treated with caution, and no general conclusions may be drawn from them, they have nevertheless been included, where appropriate, to validate or challenge findings from the case-study data.

3 A level findings

3.1 The impact of changes on teaching and learning

Evidence points to two trends that seem to be discernible since the introduction of the new specifications:

- From 1996 to 2009 (the last year in which the previous A level specifications were awarded), A level attainment continued broadly on the same improving trend in most subjects – a small steady increase of about 0.06 of a grade per year in grades attained, punctuated by short-term dips downwards from the trend at times of specification change. Also, since around 2003, participation has been increasing in most subjects.
- For results in 2010, a year in which another specification change has occurred (the specification change for teaching from 2008 working through to results in 2010), attainment saw a plateauing of grades following a long period of increasing grades – i.e. another shift downward from the trend, reminiscent of results in previous change periods in curricula (although much smaller in size than the dip following the specification change in 2000). Looking at individual grades attained (across all new specification subjects combined), the same proportion of students obtained the A grade (including those who received the newly introduced A* grade) in 2010 as received an A grade in 2009 (24.6%), whereas, for example, from 2003 to 2009, on average year on year 0.7% more students received an A grade (see Appendix 4.1).

As will be seen, these trends do not apply equally to all subjects, but the change in trend is nevertheless important. Indeed, the second of the two trends given above formed the basis of a substantial statistical investigation for A levels for this report, which looked for factors in candidates' assessment and for demographic data that might explain this change, together with any further explanations from the fieldwork undertaken. The trends have been considered in terms of changes to teaching and learning – including attainment, participation and progression – and in terms of how centres behave, and of their management of the process of change, which the case-study centres and respondents to the online survey reported on. This work includes looking at the broader context of student engagement and commitment to learning, breadth and depth of subject skills and knowledge developed, and

the generic higher-order skills encouraged and supported – such as independent learning and research skills, and the ability to conceptualise.

3.1.1 Maintaining standards over time

Ofqual noted that previous specification changes at A level have led to significant changes in grade trends and may have meant that students who took the first series of new specifications were disadvantaged. As a result, in conjunction with the A level awarding organisations in England, Ofqual introduced new requirements for awarding processes for the 2010 A level and 2009 AS level series, so as to further ensure continuity of standards:

At a time of change, it is the job of the regulators to protect the interests of candidates so that they are not disadvantaged by being the first to sit these new exams. We've worked with the awarding bodies to make sure that standards are maintained this summer, and we will continue to act in candidates' best interests over the summer.¹⁷

Awarding organisations were required to let Ofqual know, before results were published, if their proposed outcomes at grades A and E¹⁸ differed by more than a predetermined threshold from the results that would be predicted from the aggregate results for that qualification in 2009.¹⁹ The potential implications of this in 2010 are as follows:

- Awarding organisations were clearly told that students should not be disadvantaged by being the first group to take the new specifications. That is, if their year 11 GCSE grades were comparable with, or better than, those achieved by candidates in the previous year, then students should receive grades in line with the previous year – and, if this did not happen, awarding organisations were to explain to Ofqual why it had not.
- The predictive model for A level grades is based on candidates' year 11 mean GCSE grade score and an aggregate of grades in the A level subject taking

¹⁷ <http://www.ofqual.gov.uk/help-and-support/94-articles/341-changes-to-a-levels-in-summer-2010>

¹⁸ Grades A and E at A level are set by committee, with the intermediate grades set using a pre-defined formula once the A and E boundaries are set.

¹⁹ Under normal arrangements for grading, awarding organisations use a mean GCSE score of candidates' year 11 results as part of their evidence for grading decisions. In addition, for awarding in 2010, a predictor of mean grade score was used to predict A level grades, based on an analysis of the results from 2009. This predictor was created using aggregate results from all awarding organisations, whereas grading processes traditionally only consider prior awarding decisions from the awarding organisation itself. The predictor was used for grades A and E with the positioning of the intermediate grades undertaken according to the usual procedures. A similar process was used to produce a prediction model for the A* grade. The tolerance by which the award of A and E grades could vary from the model without the requirement to report to Ofqual was +/-1% (for awards with more than 1,000 candidates), and +/-2% for the A* grade. This additional information was provided as part of stakeholder discussions with Ofqual.

results from all awarding organisations – in other words, a type of 'value added' calculation for an A level subject based on GCSE achievement at the end of KS4 and grades achieved in the A level.

- Changes from six units to four and associated changes in patterns of resitting might also result in variations in the profiles of grades awarded for particular subjects by an awarding organisation, which teachers could find differed from their expectations.

Awarding meetings take account of a wide range of evidence, and the basis for weighting these in terms of final decisions is not published, so it is not possible to fully assess the impact of the changed arrangements on grade profiles.

As part of the consultation, Ofqual advised that in only a small number of awards the awarding organisations reported results outside the threshold range.

The investigations undertaken for this report confirmed the following:

- A level grades plateaued in 2010 (an increase in mean grade score of 0.17% from 2009 to 2010 compared with an average of 1.61% year on year from 2003 to 2009).
- The mean grade scores at GCSE for the 2010 A level cohort increased from 2009 broadly as for previous years.
- Demographic analysis of the 2010 A level cohort (in comparison with the 2008 cohort) offers no explanation for the plateau (see section 3.1.6).
- The threshold for awarding organisations reporting variation to Ofqual (pre-results publishing) was set at $\pm 1\%$ at A and E grades.

It should be noted that specification changes in the past have had a much larger impact on attainment outcomes.

This report shows that it is likely that the new protocol for awards affected outcomes at grades A and E (and therefore the grades in between as well), in order to ensure that candidates were not disadvantaged by being the first cohort for the new specifications. The measures introduced by Ofqual for 2010 are to be repeated (with refinements) for the 2011 A level results, noting that the purpose of the change to awarding guidance has been extended to ensure that in addition to preventing disadvantage to those in the first year of a new specification, there are also no increases in attainment in subsequent years as a result of teachers becoming more familiar with the new specifications.²⁰

²⁰ <http://www.ofqual.gov.uk/files/2011-05-16-maintaining-standards-gcses-and-alevels-summer-2011.pdf>

3.1.2 The impact of specification changes on student attainment at A2

This section compares grades achieved in the first year of the new specification A levels (2010) with teachers' expectations for these outcomes, as measured during the 2009/10 academic year leading up to the 2010 A2 examinations. As reported above, before the changes to the specifications there had been, in general, a relatively steady increase over time in the mean average grade score for A levels. Case-study centres visited in 2010, however, had predicted a change in attainment for both higher- and lower-attaining students. It was thought that, across the six subjects,²¹ without sufficient support and teaching strategies for lower-attaining students, the gap in attainment between these students and higher-attaining students might have increased, especially at A2. It should, however, be recognised that these predictions from case-study centres were based on a relatively small number of actual results at AS level and expected results for A2. Although case-study centres had already had experience of the new specification AS examinations, most had not had first-hand experience of the A2 examinations at the time of the visits. The majority of centres, across subject areas, felt that the upcoming A2 examinations in summer 2010 were less predictable. For example, many teachers interviewed had expressed concerns that they had not seen enough sample papers to fully understand what might be included in the summer 2010 examinations.

Overall grades at A2 were reported by case-study centres in 2011 to have been as expected, stayed the same, or gone up, in English literature and geography – and to have been as expected, stayed the same or gone down in French, history, physics and psychology in the 2010 assessments (but the small sample size again needs to be noted).

The key reason given for improvement in A2 grades was the fact that students were able to re-sit to improve marks. One centre, where 40% of students achieved an A grade at A2 English literature, raised the question whether it was fair for students to re-sit a module at the end of their course, in year 13, and count it for the same uniform mark scale (UMS) points as those doing it first time around in year 12. The removal of coursework from A2 geography was cited as a reason for improvement in grades, with one centre commenting that geography students seemed to do better on the content-based papers than the skills-based ones, which, they felt, were ambiguously worded.

The reasons given for the decrease in observed A2 grades varied across the subjects that reported a decline. For French, the reasons given were: the examinations were much harder; students did not learn well independently; the grades of lower-attaining students were affected by the replacement of coursework with a written essay examination. One

²¹ In 2010, media studies was included in these subjects; its place has now been taken by history.

centre, commenting on history, stated that the style of questioning was very academic, which resulted in a widening of the gap between the higher- and lower-attaining students. Another centre reported that the structure of the physics exam was hard to follow for students, who found the contextual use of physics, on which there is now more emphasis, difficult and therefore did not always manage to finish in the allotted time. The removal of coursework as part of the assessment was cited as a reason for the decrease in A2 grades in psychology.

Table 1 gives the results in these subjects for the whole year 13 candidature at A level; the table tells a rather different story from the one told by case-study centres.

Table 1: Case-study schools' expectations of 2010 A level results compared with actual results for the whole candidature

Subject at A level	Expected/ observed grades at case-study centres	Mean grade score result for 2010 candidatures as a whole (SFR)
English literature	Stay the same or go up	Decrease of 0.01 grade points compared with 2009, following an average increase of 0.04 each year from 2003 to 2009. However SFR data does not disaggregate for English literature, so this should be used with caution
Geography	Stay the same or go up	Increase of 0.01 grade points compared with 2009, following an average increase of 0.06 each year from 2003 to 2009
French	Stay the same or go down	Increase of 0.05 grade points compared with 2009, following an average increase of 0.06 each year from 2003 to 2009
History	Stay the same or go down	Increase of 0.03 grade points compared with 2009, following an average increase of 0.05 each year from 2003 to 2009
Physics	Stay the same or go down	Increase of 0.05 grade points compared with 2009, following an average increase of 0.04 each year from 2003 to 2009
Psychology	Stay the same or go down	Increase of 0.01 grade points compared

Subject at A level	Expected/ observed grades at case-study centres	Mean grade score result for 2010 candidatures as a whole (SFR)
		with 2009, following an average increase of 0.05 each year from 2003 to 2009
All new specification subjects		Increase of 0.01 grade points (<i>the plateauing described earlier</i>) compared with 2009, following an average increase of 0.05 each year from 2003 to 2009

Teachers' expectations in the case-study schools do not reflect the results for the candidature as a whole, which suggests that either their schools are misrepresentative of the whole or (more likely, given that the case-study schools are a balanced sample) the actual results in 2010 were less predictable for teachers than in previous years, as has also been the case at the time of previous specification changes.

The mean grade score aggregates the proportions of candidates achieving each grade. Appendix 4.1.2 looks at the proportions of candidates achieving each grade for the new specification A levels in 2010 and preceding years. It shows that in 2010 changes in the proportions achieving each grade differed from the trends in preceding years: the proportion of A/A*²² and B grades plateaued following steady increases since 2001, the proportion of C grades increased a little, D grades plateaued after falling since 2001, and E grades continued to fall.

Within individual subjects, changes to trends in proportions achieving particular grades were also seen, generally reflecting the overall trend but with some differences (see Appendix 4.2). Comparing 2010 results with those in 2009, the proportion of A/A* and C grades fell in English, while D grades increased. In geography, the proportion of A/A* grades fell, while B and C grades increased. In French, the proportion of B and C grades increased, while A/A* grades plateaued. The proportion of A/A* grades increased in history in 2010 compared with 2009 (the only one of the six subjects for which this was the case), while B and C grades plateaued. In physics, the proportion of A/A* and C grades plateaued, while B grades increased; and the proportion of A/A* grades fell in psychology, while B grades increased. Overall, there were more changes in the proportions of candidates achieving each grade in 2010 than would have been predicted from previous years' results.

²² For all year-on-year comparisons, the A* grade is considered equivalent to an A grade.

There were limited reports on any changes to AS grades. Some centres had requested copies of students' examination papers and subsequently requested remarking by awarding organisations. One centre reported inconsistent marking in 2009 on AS English literature but also said that the grades for A2 were in line with what they had expected, suggesting that the differences had been evened out in the awarding process. A second centre reported inconsistencies in marking on both AS and A2 French assessments, and a third on AS and A2 history.

The online centre survey provided some support for the view that it was more difficult for students to achieve their expected grade at AS level since the changes in the specifications, with 24 centres saying that there was no change, 20 that it was harder, and only 3 that it was easier. (There was least support for the idea that it is harder for students to achieve their expected grade in English literature.) There was similar support for the same view at A2, with 23 centres saying it was harder, 16 that there was no change, and only 3 that it was easier.

The online centre survey provided some data on student grades in 2009 and 2010, both from senior leaders and from the individual subject surveys (the latter providing data at A2 as well as AS). For AS level, the survey reported fewer A and B grades in 2010 compared with 2009, with more C and, particularly, D grades.

The most noticeable aspect of the A2 grades reported in the survey is that the percentage of students graded at A has decreased, as might be expected, as a proportion of them now achieve A*. However, there are also more students reported as achieving grade C, which is not a grade that is awarded judgementally²³ by examiners. The online survey also captured a small number of comments from teachers relating to AS and A2 unit results. A number mentioned fewer A grades, and there were specific comments about the difficulty of achieving an A grade in speaking in French.

3.1.3 Introducing stretch and challenge

'Stretch and challenge' is being interpreted in a variety of ways, and the degree of uncertainty and ambiguity about what is involved is reflected in the responses from both the case-study centres and the online survey.

QCDA defines stretch and challenge in the teaching and learning context as: developing higher-order skills of analysis, evaluation and synthesis; advancing conceptual

²³ At A level, senior examiners set grade boundaries at A and E using a combination of professional judgment and statistical evidence. These are known as 'judgemental' boundaries. Boundary marks for the other grades are calculated arithmetically.

understanding and linkage between concepts across the specification; and improving the quality of written communication.²⁴

*Stretch and challenge places renewed emphasis on developing conceptual understanding while maintaining breadth of knowledge. Also, stretch and challenge involves developing ‘an understanding that ties together and makes sense out of this information [...] and also the ability to communicate that understanding’. Moreover, ‘students need to be supported to recognise and understand key concepts and ideas, and then relate and transfer them to a range of different contexts’.*²⁵

The guidance further notes that ‘achievement on stretch and challenge will be recognised, rewarded and reported through an additional A* grade’.²⁶

The definition of stretch and challenge given by QCDA indicates that it should be built into the whole teaching and learning fabric of A level courses, to raise the levels of achievement of all students. This ambition, however, appears to be at odds with the statement that it is recognised only in the case of those who gain A* grades.

Similar uncertainty about whether the stretch and challenge requirement is directed at all students, or only the highest attaining, is apparent elsewhere. An A level briefing from Ofqual, for example, maintains that:

*The new specifications were designed to include, in the A2 papers, greater ‘stretch and challenge’ for the very top candidates ... So some of the questions on the new A2 units will be tough, because they are designed to challenge the very best candidates.*²⁷

However, the monitoring report on the 2010 examinations suggests that stretch and challenge has a wider focus, finding that:

*In general ... the A2 assessments contained an appropriate degree of stretch and challenge ... thereby ensuring that they gave scope for the full range of candidates – including the most able – to demonstrate and be credited for what they knew, understood and could do.*²⁸

²⁴ QCDA (2010) Stretch and challenge and the A* grade: Guidance on changes to A level teaching and learning, p7.

²⁵ QCDA (2010) Stretch and challenge and the A* grade: Guidance on changes to A level teaching and learning, p8.

²⁶ QCDA (2010) Stretch and challenge and the A* grade: Guidance on changes to A level teaching and learning, p5.

²⁷ Presentation at an A level briefing event, 17 May 2010, Ofqual.

²⁸ The New GCE A Level Examinations: Findings from the Monitoring of the New Qualifications in 2010, Ofqual, April 2011, p19.

In terms of assessment, the criteria require that arrangements in GCEs must include A2 assessments that provide appropriate demand through:

- the use of a variety of stems in questions
- ensuring connectivity between sections of questions
- requiring extended writing in all subjects, except where agreed as inappropriate with the regulatory authorities
- using a wide range of question type to address different skills
- including synoptic assessment.

Awarding organisations are, therefore, introducing stretch and challenge in one or more of a number of ways:

- into the qualification syllabus (e.g. increased content, higher-order skills/concepts)
- changes to the assessment objectives and/or relative weightings
- in the paper setting, via amendments to the nature of the questions/tasks
- in the marking criteria
- at the grade-awarding process.

The different approaches depend in part on the way skills and knowledge are organised within a subject. For example, in modern foreign languages knowledge, understanding and skills are closely linked, and synoptic assessment promotes stretch and challenge.

One awarding organisation's AS/A level history specification notes that:

*The requirement that stretch and challenge is included at A2 is met... by means of... questions [that] will enable candidates of differing abilities to be challenged at their own level and allow the most able to show the breadth and depth of their knowledge and understanding. The requirement is also met by requiring candidates to complete a historical enquiry... The descriptor for the highest levels of response in the mark scheme will ensure that the most able are stretched...*²⁹

Some of these approaches (e.g. amendments to the content or question paper) are more readily perceived than others, and centres may then re-focus their teaching and learning appropriately. The case-study data reports centres interpreting stretch and challenge in terms of developing independent learning and research skills; synthesis and conceptualisation of ideas across the subject; depth and breadth of learning and the A*

²⁹ GCE AS and A level Specification, History, AQA, 2007.

grade. Particular tensions may arise when stretch and challenge in the syllabus, and hence in the teaching and learning, is not matched, or is not perceived to be matched, by the styles of questioning and marking strategies employed in the assessment.

3.1.4 The development of subject-specific skills and knowledge

Data from the visits to the case-study centres in spring/summer 2011 reported a consensus across the six A level subjects that the new specifications had brought about positive changes to teaching and learning. In French, for example, it was felt that the removal of coursework had helped students work on topics in greater depth, which was a new requirement. In particular, the reduction from six units to four (in all subjects except physics) was seen, not only in French but also in English literature and geography, as having introduced greater challenge and depth into teaching and learning, encouraging both the development in depth of subject-specific skills and greater breadth of subject-specific knowledge.

These recent findings were not wholly consistent with those from previous visits. In summer 2010, case-study centre teachers of different subject areas and those working with different awarding organisations had appeared to disagree about the impact of the reduction in the number of units. There was some concern, for example, that an increase in the amount of subject content could lead to teaching and learning being so overloaded by the need to input content that the development of higher-order subject-specific and generic skills would suffer: depth might be sacrificed for volume. And, in a few case-study centres, teachers reported that they continued to use didactic, essentially transmission-based, rather than developmentally based, approaches to teaching because they felt there was so much subject content for students to be given.

Similarly emerging from the initial case-study visits was a divergence of view about the merits of the greater emphasis in the specifications on independent learning. This was generally seen as positive and something that stretched and challenged students by encouraging them to work in more exploratory and autonomous ways – particularly where the course content was seen to be no longer content driven, but skills driven. In such instances, teachers felt that a new emphasis on activities like decoding research, looking at more abstract problems, dealing with different conceptual issues and viewpoints – as opposed to learning facts and regurgitating them – had all been positive, albeit challenging, changes to the specifications.

Conversely, some saw drawbacks to this new focus on independent learning. A few teachers commented that a transition period was needed while students got used to working in a more independent way. There were some concerns that the overloaded course content at

times meant that independent study was very outcome-focused rather than truly independent. It was also felt that the focus on independent learning was carried into neither the examination questions nor the mark schemes, which did not always ask for, or credit, independent thought and investigation.

In 2010, many teachers had felt that the fact that there was insufficient time for students to respond in an exploratory way in the examinations was not in the spirit of the new specifications, which sought to promote a more exploratory approach, to reward originality of thought, and to give less emphasis to the regurgitation of facts. Examinations that were narrow in style and choice of question failed to reflect this, with the exception of geography, in which teachers reported that the more open-style exam questions at A2 did allow for greater originality, individuality of thought and creativity.

These points were picked up again in the follow-up visits. Teachers of English literature, French and history, as well as geography, now thought that the new assessment regime was encouraging independent thinking and in-depth learning. Teachers of physics and psychology, however, remained unconvinced, continuing to feel that the assessment regime failed to achieve the aim of stretching and challenging students. There are issues here that would warrant further exploration, such as whether assessment in science subjects tends to employ fewer open questions than humanities subjects and, if so, whether closed questioning is somehow less amenable to the notion of stretch and challenge.

From the online centre survey of individual subjects, of the 59 respondents who expressed a view on the extent of stretch and challenge at AS level:

- 29 thought that students were equally as stretched and challenged as previously
- 24 thought that the changes to the A level specifications had stretched and challenged their students more
- 6 thought that there was less stretch and challenge.

The subject most often perceived as exhibiting greater stretch and challenge than previously was geography, with physics the least.

The online survey also showed modest support for the idea that new A level specifications created greater stretch and challenge at A2. Of the 43 who expressed a view, 28 thought that the assessment created greater stretch and challenge, and 25 thought that the content did. Again, the subject for which this view was expressed the least was physics.

3.1.5 Opportunities for higher-level generic skills development

During the visits in 2010, the ability to engage in independent and investigative learning, and the capacity to think holistically, were commonly cited as examples of higher-level generic

skills. These skills were seen as being not only valuable in a subject-specific context, but also significant in enabling entry into higher education. Opportunities for their development in the specifications were felt to be provided by:

- greater flexibility in choosing course content
- more-engaging choices of topic or areas of thematic study
- greater emphasis on the application of knowledge and skills
- the synoptic units.

The follow-up visits to the case-study centres in 2011 found general agreement that the new English literature and geography specifications and assessments encouraged the development of synoptic skills; a minority of centres considered this to be the case for history, too. However, about half of the respondents who were asked about English literature and geography thought that it was too ambitious to try to develop these skills with this age group and that they would be more suited for first-year undergraduate study.

Teachers from three subjects – English literature, geography and psychology – expressed the belief that the changes to the specifications had led to a greater focus on the application of skills and knowledge. Although this was a challenge, these teachers felt that it was a desirable move, as it provided a better preparation for study at university. In AS and A2 physics, however, teachers felt the examinations were now easier and less demanding: for example, students were no longer required to learn key equations.

About half the centres commenting on English literature and history thought that coursework was helping to develop greater higher-order skills at A2, whereas the external assessments at AS were felt to be hindering such skills development. Most respondents commenting on English literature and half of those commenting on history felt that the coursework element of the new A2 specifications encouraged independent thinking and the development of research and writing skills. Many teachers considered coursework a useful tool to motivate and reinvigorate students and to stretch and challenge them through differentiated work, for example by adjusting the focus of their role from transmitter to facilitator and tutor. They also felt that coursework allowed them to provide more-individualised support and structuring to help lower-attaining students get marks ‘under their belts’ before the examinations, while also challenging higher-attaining students to extend and develop their skills fully.

The final unit of A2 history and French was also seen to encourage independent learning. Most centres that discussed geography, however, thought that the independent learning and thinking required of the students in the new specifications was too challenging for them.

The majority of A level students' subject- and non-subject-specific comments suggested that they would wish to retain coursework as part of the assessment regime; that coursework performance was a better indication of a learner's ability; that it removed some of the pressure brought about by examinations; and that it allowed students to show different skills from those evidenced by writing under examination conditions. Physics students suggested that physics did not lend itself to coursework, but that other subjects did.

The extended project qualification was considered by several centres to offer opportunities for students to develop higher-level generic skills.

3.1.6 Participation and student engagement

Participation for the cohort completing new specification A levels in summer 2010 followed the upward trend of recent years in the subjects considered here – except for French, which continued with no increase in numbers (an average decline of 1.2% year on year since 2003), and psychology, where numbers grew sharply (an increase of 7% on 2009) after several years of little growth for males and females (see Appendix 4.3 for subject breakdown).³⁰ The same is found when considering the 2010 AS level results, which similarly continued the same upward trend of recent years except for 2008, when there had been a small dip, before the specification change. However, it should be noted that the AS level candidature is known to under-report the actual number of candidates taking AS level qualifications – as some candidates record no qualification in year 12 because they do not cash in their AS level results until year 13, at the time they submit for the A level qualification. Official Statistics and NPD data record only the 'highest' qualification in a subject for a candidate in any given year, so when a candidate cashes in their AS results at the same time as their A level qualification (rather than having cashed in in the previous year), the AS award is not recorded at all in Official Statistics, either in year 11 or 12 results. The extent of this under-reporting has not yet been fully investigated but was, for example, 25% for AS history in 2010. As a result, SFR data cannot be used to accurately report progression from AS to A level (in the history example cited, SFR data would suggest progression at 96%, whereas the actual figure is 73%).

Centres make decisions about when to cash in units (and AS cash-in delays are more common in independent schools, for example). These decisions can be influenced by a number of factors, including plans for re-sits, school performance tables and students' own attainment evidence for UCAS (i.e. it may be seen as better to show no AS outcome than a poorer grade than the A level grade attained – although some universities ask for complete

³⁰ A level mathematics and A level further mathematics, for which specifications at A level have not changed, were not included.

unit level information as part of the application process). With the reduction from six units to four, and the introduction of the A* grade (to which AS units do not contribute), it seems that cashing-in behaviour may change with the introduction of new specification A levels – the final report will investigate this further. From 2011 on, changes to funding arrangements for maintained schools mean that schools may lose funding for candidates who do not certificate their AS achievements at the end of year 12; this is likely to have an impact on future cash-in strategy.

The possibility that other demographic variation in the cohorts could offer an explanation for the change in grade profiles achieved at A level in 2010 was considered. An inspection of candidature variations by the key demographics for the 2010 A level candidature (shown in Appendix 4.4) suggests little difference to the preceding year (although this is an area that would benefit from further research in subsequent stages): the overall A level candidature for the new specification subjects in 2010, in other words, is largely as trends would suggest. Analysis of the composition of the 2010 year 13 cohort shows little unexpected change from the previous year – proportions of students from deprived backgrounds³¹ and those with English not as their first language are broadly as in 2009, and the mean grade score achieved at GCSE for candidates in 2010 was higher than for those in 2009 (reflecting the trend of increasing grades in recent years). The observed increase in numbers of SEN students taking A levels in 2010 may be related (in part at least) to increases in the number of SEN students both in KS4 cohorts and progressing to KS5. Small exceptions are noted for:

- An increase in the proportion of students taking geography A level who do not speak English as their first language.
- A reduction in the proportion of students taking history A level who were eligible for free school meals (FSM) in KS4.
- An increase in the proportion of French A level candidates who were eligible for free school meals in KS4.

Appendix 4.4.2 shows the prior GCSE attainment of candidates completing A levels in 2008, 2009 and 2010. Generally, candidates completing A levels in 2010 had better GCSE grades than their predecessors. In conclusion, therefore, the cohort completing A levels in 2010 is broadly as would be expected, based on previous years, in terms of their demographics and prior attainment.

³¹ The demographic variables shown here are not recorded consistently in the KS5 NPD data so these demographic variables are derived from matching the KS5 data to the relevant year's KS4 census data. Note also that, for Special Educational Needs (SEN) calculations are based on the proportion of candidates that had any of the three types of SEN provision at KS4.

3.1.7 Changes in participation

Early indications within the case-study centres in summer 2010 suggested there had been some fluctuation in participation both in terms of numbers and student profile (for example, according to prior attainment or as a result of including students who may have previously followed a vocational course). It was difficult to tell, however, whether this was directly linked to changes in specification or to other school-based reasons. A minority of case-study centres had already raised the criteria required for entry to A level psychology and English literature courses, as they had found that students with less than a B grade at GCSE struggled with the transition to AS. In a number of centres students were taking an increased number of AS courses in order to gain a broader education – and delaying specialising further until after the AS results.

Case-study centres reported that increasingly students were ‘playing it safe’ by dropping subjects in which they had not achieved the desired grade in the AS examination and opting for the subjects that they considered ‘easy’. What students define as ‘easy’, however, often reflects their own earlier experience of the subject itself or their experience of particular teachers; there are not necessarily particular categories of subject that are considered ‘soft’ options. Feedback on courses from other students is also a factor when students are choosing their A2 subjects.

In 2009 nationally, the numbers taking French had fallen significantly, with 50% fewer students taking the subject at A level than took it in 1996. In the case-study centres visits in 2010, numbers had either gone down or stayed as they had been before the specification changes. The decrease in AS participation was attributed to three factors: students finding GCSE French very challenging; A level students telling GCSE students how difficult A level French was; and more-able linguists opting for the International Baccalaureate (IB) rather than A level. In the more recent follow-up visits, three of the case-study centres reported that students appeared to be more engaged with new topic areas and genre choices in French. This view was reinforced by students from one centre who spoke of a high level of engagement with the subject.

One case-study centre reported that they had a significantly higher-than-usual rate of history AS students who did not intend to carry on to A2, because they felt that too much had been crammed into the history AS course and there was insufficient variety. Similarly, during the case-study visits in 2010, it was reported that there were fears that fewer students would continue with geography after AS because the course was felt to have become more challenging.

Based on data from the online centre survey for individual subjects, however, there is no compelling evidence to suggest changes in the pattern of students choosing to go on and study a subject at A2.

3.1.8 Preparation for progression

3.1.8.1 Progression from GCSE

In addition to the demographic information about participation, this report considers whether the changes in attainment trends may result from changes to the profile of ‘incoming’ students to the A level programme: for example, if the students starting the new specification A levels (in 2008, completing in 2010) were of a weaker profile than the equivalent group starting the last cycle of the old specification A levels (in 2007, completing in 2009), this might offer an explanation for the plateau in grade trends observed.

The investigation was undertaken first by comparing A level outcome grades for students with the grades they achieved at GCSE in the same subject (e.g. comparing physics GCSE grades with physics A level grades). No clear pattern emerged from this comparison, suggesting that, for these subjects at least, changes in A level grades from 2009 to 2010 are unrelated to any changes in associated GCSE scores from 2007 to 2010.³² This finding also supports the view (presented in the preceding section on participation) that the students taking new specification A levels are very similar in profile to those taking the old specifications immediately previously.

The above analysis considers progression in terms of a single GCSE. With many candidates achieving A/A* in subjects at GCSE that they go on to study at A level, the observation of any change is limited by the range of grades which the majority of candidates achieve. The analysis also considered the effects of candidates’ mean GCSE score³³ across all GCSEs attempted, based on the ABDA dataset. When the mean GCSE score for candidates completing their A levels in 2008 and 2010 was compared, the 2010 A level candidates in each of the five subjects (English literature, French, geography, psychology and physics – see Appendix 4.4.2) obtained higher mean GCSE grade scores than their equivalent A level candidates in 2008 – so their results at A level would have been expected to show an improvement on previous years, rather than the plateauing that was observed.

³² Many schools expect students who wish to study a subject at A level to have an A*, A or B grade in the equivalent subject at GCSE (requirements may be higher in some subjects). With most students having one of the three top grades at GCSE, it could perhaps be expected that little change would be seen here.

³³ Mean grade score combines proportions of candidates achieving each grade in a subject into a single grade score for the subject. The higher the grade score, the better the candidates did overall. It is calculated by assigning a weight to each grade and multiplying that weight by the number of students achieving the grade. For example, A/AS level calculations A and A* grades are weighted as 5, B as 4, C as 3, D as 2, and E as 1. Different weightings are used for GCSE.

3.1.8.2 Progression from AS to A2

In 2010 for half the subjects, case-study centres reported that there had been fewer students progressing from AS to A2 since the new specifications were introduced. The decrease had been most dramatic in psychology, where the introduction of a science focus had led to only about half the student cohort continuing after AS. French and geography had also seen a reduction in the number of students progressing from AS to A2. There had been no significant changes in the other three subjects.³⁴ English literature case-study centres reported no particular effect of the changes on the rates of students progressing from AS to A2, noting that these rates continued to be as healthy as ever.

During the follow-up visits in 2011, only the French AS was seen as more demanding than before and leading to improved preparation for A2. The level of difficulty of the reading texts at AS was reportedly higher than those at A2 French. Two centres also suggested that the standard expected of students in GCSE French ought to be raised, to prepare students better for the new AS French.

For English literature, history and geography, those who responded thought that the AS was preparing students less well for A2 than before. The reasons given were: lack of subject-specific skills development (English literature), lack of variety and too much crammed in (history), and lack of emphasis on essay-writing skills (geography).

3.1.8.3 Progression from A level to further study

As to whether there have been changes that would have improved students' preparation for progression to higher education study, of the eight responses given by the case-study centres (across five subjects) five were positive and three negative. On the positive side, A level English literature and A level French were reported to prepare students better for higher education study, because of the enhanced emphasis on synoptic learning and a greater focus on essay-writing skills, respectively. On the negative side, A level geography was seen to focus too much on geographical knowledge instead of the research skills that universities are interested in.

Respondents to the online centre survey for individual subjects in summer 2011 indicated that there was little change in the number of students choosing to continue with their subject after A levels: 35 reported no change, 18 a slight increase, 4 a significant increase, and 4 a slight decrease.

³⁴ History was not included in 2010.

There was fairly strong support from teachers in the online centre survey for the contention that the new A* grade had fulfilled its purpose of rewarding the highest attainers, with 40 out of 53 who expressed a view agreeing, against 13 who disagreed. However, only 14 respondents reported that any of their students had been offered university places which required A* grades, with 6 of these being from respondents to the history questionnaire. Among respondents to the senior leader survey, who perhaps have more of a whole-school perspective, 11 out of 16 reported that some of their students had been offered university places which required A* grades in summer 2010.

Most of the students' comments on the impact of the A* grade were about how university places would be affected by the introduction of the A* grade. These comments did not say it would be harder or easier to gain a place, but that a university would be able to identify more appropriate students or that universities might change their entry requirements.

3.2 The impact of change on centre behaviour

3.2.1 Choice of specification

Although not always stated explicitly, the three main reasons for a centre's choice of one specification rather than another were: that it best met the needs and aspirations of their students; that it was the one that was most consistent with how they viewed the nature of the subject; and familiarity – often long-standing – with a particular awarding organisation. These reasons were uppermost in the responses across the range of subjects, but there were, within them, a number of more-specific considerations.

3.2.1.1 The needs and aspirations of students

The concern that the specification should be accessible and student-friendly was expressed by three of the centres. The characteristics of accessibility and student-friendliness were seen in terms of:

- course content – for example in geography, where students undertook a local area study
- course structure – for instance in modern foreign languages, where incremental assessment was seen as being fairer on students; where coursework, in history, offered more variety or, in English, allowed students an element of independent choice; and where a centre welcomed the existence of the synoptic unit in English
- the perceived interest of the course for students, in relation to modern foreign languages and history

-
- the course's appropriateness for students with a range of attainment in a given subject (the reference here was to physics).

3.2.1.2 Consistency with the nature of the subject

In terms of consistency with the subject's nature, the particular features that were referred to were:

- studying a particular period rather than random texts in English
- in modern foreign languages, the linguistics-based nature of the course and the choice of topics in French
- the requirements of the practical component in physics
- a more issues-based specification in geography.

3.2.1.3 Familiarity with the awarding organisation

Four of the centres responded that they were remaining with the same awarding organisation as previously on the grounds of familiarity. Among other advantages it was felt that this would make administration simpler. Only in two cases were centres either changing, or contemplating changing, the awarding organisation because of concerns about the new specification on offer.

Allied to familiarity with the awarding organisation were considerations about teachers' expertise and experience. Four of the centres referred to this: that the specification allowed teachers to follow their interests, that it was the better option because it related more readily to teachers' specialisms and, in one instance, that the teacher was a former examiner.

3.2.2 Management and resource implications for centres

There appeared to have been little impact on management time as a result of the changes to the A level specifications. Where centres reported differences, they tended to focus on the impact of coursework on workload rather than the management of A level. Any increase in the resource burden was also seen as adding to the workload of subject teachers. One centre, for example, felt that – as a result of cuts in funding and the reduction in examination officer time – teaching staff were taking on more of the burden of examination administration, which added to their workload. In the online centre survey, managing and structuring the whole scheme of work for the A level course was seen as the single most significant change that respondents reported seeing since the introduction of the new A level specifications.

A number of respondents to the online survey reported implementing greater use of IT and/or online resources when moving to the new specifications. In particular, respondents in

modern foreign languages spoke about the greater use of MP3 players and computers in learning and in examinations.

3.2.2.1 Management of time

Centres commented on the impact on teaching time primarily in relation to the changes made to the examination timetable. One head of physics found the changes to both the timing of AS and A2 problematic: the AS examination was considered to be too early, leaving less time for teaching, while the A2 examination came very late. The head of department's concern was that the second-year students had six weeks between going on study leave and taking their examination, again losing valuable teaching time. The same centre also expressed concern at the turnaround of results from January examinations, which left little time to re-enter students for re-sits and also made it difficult to assess students' progress. In the online centre survey, respondents reported that approaches to classroom teaching and activities were the second-most significant change since the introduction of the new A level specifications.

For the ten centres in the online centre survey offering the extended project qualification, half reported that students start work for the extended project in their AS year, the other half that they start work for it in their A2 year. All reported that their students are assessed on their work for the extended project qualification in their A2 year.

3.2.2.2 Re-sits

A number of centres commented on the number and frequency of re-sits. In four centres, examination officers reported an increase in the number of re-sits for A level. One examination officer reported that in some subjects one-half to three-quarters of students were re-sitting examinations, while another described the situation in terms of 'becoming a re-sit society', in which it was fast becoming the norm to re-sit if the grade was not high enough. As one examination officer commented, however, this was not so much to do with the change in specifications as with an increase in student numbers, together with the pressure on students to gain as high a grade as possible. Nonetheless, the increase in re-sits was resulting in additional work for the examination officer.

Students welcomed the opportunity for re-sits, with the majority believing re-sits should be allowed in any circumstance and only a minority suggesting that re-sits should be allowed in only exceptional circumstances.

Another possible explanation for the plateau effect seen in the attainment data is a change in re-sitting behaviour or the outcomes of the re-sits. The 2010 ABDA data lacks the wealth of information about re-sits of the 2008 data, but some insight can be gained by examining

the change in UMS scores between re-sits of unit 1 at series 1 (i.e. the first attempt) and series 2 (the first re-sit).³⁵ Re-sits of unit 1 increased from 7.1% of the candidature in 2008 to 9.6% in 2010.³⁶ If re-sitting leads to an improvement in grades and with more re-sitting in 2010, one might expect UMS scores (and grades) on average to have increased more between the first sitting and the re-sit in 2010 than in 2008. This is not found to be the case. In both years the effect of re-sitting is small (an average gain across the re-sitters of 7 UMS points for the unit), but the gain is smaller in 2010: an average of 5 UMS points for the four subjects where an improvement is seen: English, geography, psychology and physics. Only for geography is the average re-sit improvement higher in 2010 (5.7 UMS points) than in 2008 (4.1 UMS points). The 58 re-sitters of unit 1 in French in 2010 actually performed less well than on average than in their first sitting. Unfortunately, analysis of trends across a wider range of years is not possible, as re-sit data is only available for 2008 and 2010.

The evidence suggests, however (even if based on this crude and limited analysis), that changes in re-sitting behaviour and in the outcomes of re-sitting are unlikely to be driving the plateau effect seen in the grade distributions between the two years.

3.2.2.3 Staff development

The majority of responses relating to staff development focused on the sessions offered by awarding organisations. Teachers were largely positive about their experiences of training from awarding organisations, some citing the opportunity to meet colleagues as an additional benefit. One centre said that the bespoke training, where the awarding organisation came to the school, was of the most benefit to them; another liked the online training offered by one awarding organisation, not least because it was not too expensive. The same teacher felt that going through the training more than once helped them to gain confidence in 'how to stretch (students) to get the synoptic bits'. Another teacher of English, while welcoming the training on stretch and challenge, would have welcomed more help with stretching the lower-attaining student rather than always focusing on the higher-attaining student. The need to attend training more than once was echoed by a history teacher. Another head of modern foreign languages agreed that awarding organisation training would be a useful way to learn about written coursework and to meet colleagues. In the online centre survey, respondents also mentioned the help and resources provided by awarding organisations. There were both positive and negative comments, but it is clear that teachers have relied quite heavily on the materials, resources and training provided by the awarding organisations to come to terms with the new structure of the qualifications.

³⁵ A level history is not included in the ABDA data.

³⁶ The years here refer to the year of completion of the A level.

Other examples of training which staff found useful were those that were internal to the institution. One centre had introduced a coaching model to share knowledge between teachers. A head of physics felt that 'training on pedagogy and the constructivist approach would be [more] useful' than focusing on content, going on to say that they attended updates at the Institute of Physics. The same head made the point that the introduction of the 'rarely cover' policy made attending training in the school week more difficult. One English teacher would have liked more help with using a virtual learning environment (VLE) more effectively, although they felt that less-experienced teachers might welcome guidance on aspects of wider reading. One centre was looking at differentiation between students in lessons rather than focusing on marking work outside lesson time.

3.2.2.4 Impact of changes on staff workload at centres

Teachers' perceptions of the increase or otherwise in their workload varied from subject to subject. One head of French said that the reduction from six units to four had decreased their administrative workload. One head of English thought that some of the increased workload was due to the need to become familiar with the new specification, but then went on to say that the coursework element had added to their workload in terms of preparation, teaching and assessment. One head of history agreed that the workload had increased with the introduction of coursework, while another said that the requirement for two pieces of coursework was over-burdensome for their staff. They went on to explain that some of the pressure was due to the need for students to study a considerable amount of the course before undertaking coursework; this had the effect of squeezing two pieces of coursework into a short time-span, and the knock-on effect of teachers needing to turn round the marking and moderation very quickly. However, a head of geography felt that the introduction of the new specification had made little difference to the workload.

Just under half of the case-study centres currently offered the extended project. One head of psychology talked about how much they had enjoyed working with students on the extended project despite the increase in workload, in part because teachers were not necessarily supervising a project within their own discipline and so 'you do have to do a bit of reading and inform yourself (as well as) keep(ing) on top of the paperwork'. A member of staff at the same centre echoed this, saying the extended project was a source of enjoyment and an opportunity for their own development and enrichment.

Ten out of sixteen respondents to the senior leaders' survey stated that their students are offered the extended project qualification. Four of these centres stated that they had an extended project co-ordinator, and a further three that co-ordination is the responsibility of the director or head of sixth form. Four centres stated that the extended project co-ordinator

(possibly in conjunction with others) delivers the taught element of the qualification; the remaining centres use a variety of other staff, including librarians, to deliver the taught element. Similarly, it is the responsibility of the dedicated co-ordinator and/or supervisory team (where there is one) to co-ordinate and organise the monitoring and support of students' work, while for other centres the work is carried out by teachers, tutors, etc. Ten out of sixteen senior leader survey respondents reported that they had seen evidence that universities are taking the extended project qualification into account when making offers to their students.

4 GCSE findings

4.1 The impact of changes on teaching and learning

Further statistical analysis on GCSE data will be undertaken for the final report once the new specification GCSEs are awarded in summer 2011. The baseline data analysis for overall grades in all GCSE subjects up to now has shown a steady increase over time.

4.1.1 The development of subject-specific skills and knowledge

Five case-study centres discussed the respective merits of unitised and linear assessment in GCSE French, history and mathematics. They were of the view that linear assessment promoted more in-depth and longer-lasting learning than unitised assessment, particularly in relation to the development of subject-specific skills. Two of the other centres stated that the amount of content they had to cover, allied to the drive to prepare students for examinations, hindered the development of in-depth subject knowledge.

Two centres thought that the new speaking test in GCSE French and Spanish enhanced the development of subject-specific skills by guarding against students simply regurgitating pre-learnt material. It was also noted, however, that the test expected a degree of spontaneity beyond GCSE level. Some modern foreign languages teachers felt that the reduced emphasis on grammar skills and on 'passive skills' (listening and reading) in the assessments disadvantaged lower-attaining students.

The general view was that the new specifications did not encourage the development of subject-specific skills in GCSE geography and history. In geography, the reason given was that covering all physical geography topics in one year and human geography topics in another year failed to engage students and did not help them to develop their data response skills – something to which the physical modules do not lend themselves well. In history, the reasons cited were that the course was too easy overall and that the unitised model of assessment did not give students enough time to develop their skills.

In the online centre survey, nearly half (23) of the respondents did not feel that students were able to gain a deeper, more-lasting understanding of the subject as a result of the changes to the GCSE specifications, and 17 felt that there was no change. Only 11 respondents (English, French, geography and mathematics teachers) felt that students did gain in this way. No respondents from history or Spanish agreed.

In the senior leaders' online survey, there was a fairly even spread among the respondents between those who had opted overall for linear assessment, those who had opted for unitised assessment, and those who had opted for a combination of both.

Approximately half of those providing an explanation indicated that the choice of unitised or linear assessment is a pragmatic one that depends mainly on the subject. However, respondents had varying and sometimes diverging views on the merits of unitisation. Respondents' concerns included a belief that unitisation destroys educational coherence within subjects, and does not encourage students to see how the parts 'fit together'. There was also a view that unitised examinations (and re-sits) may improve results, but do not help with a stronger understanding of the subject – too much time is spent examining, and not enough teaching and learning. However, there was also support for unitised assessments, which can be seen to allow students to achieve along the way – and to re-sit if necessary.

The students, however, were in favour of the unitised approach for a number of reasons: it made learning easier (as it was presented in bite-size chunks), the content was clearer, and the material was fresher in the learner's mind when it came to the examination. Students also liked the fact that unitisation meant that there were exams throughout the year rather than a single exam at the end. They felt not only that this took away some of the stress of exams, but also that it was easier to revise and it made all the topics link together better, as they learnt a whole topic in one section before moving on.

4.1.2 Opportunities for higher-level generic skills development

Three case-study centres, commenting on GCSE geography, expressed the view that the unitised approach to assessment in GCSE geography does not encourage synoptic learning. These centres did not see this as a major problem, however, because teachers make the connections between the themes for students.

Five centres discussed the style of questioning and assessment criteria for the functional and problem-solving questions in GCSE mathematics. While it was thought that these promote independent thinking, as students need to think through a problem and break it down into stages, there were concerns over the level of literacy that students needed to tackle functional and problem-solving questions.

Centres did not discuss higher-level skills development extensively for GCSE English. One centre reported that controlled assessments³⁷ promote independent thought and learning, and another centre that there has been a decreased emphasis on analytical skills in favour of summarising skills.

4.1.3 Participation and student engagement

Most staff interviewed at the case-study centres considered that students were equally or more engaged and motivated than before, but contrasting views were expressed for modern foreign languages, geography and mathematics.

Controlled assessment was seen to have made students of French and Spanish more focused, and learning was thought to be more interesting overall, although there appeared to be particular issues with the new speaking assessment.

Against this, however, one case-study centre and several centres responding to the online survey reported that ‘the fun had gone out’ of the teaching because they no longer had time to pursue things that are not part of the assessment. They took the view that considering content and/or skills that were unlikely to be assessed was an indulgence. Students were considered to be very outcome-focused and aware of what qualifications they need for progression. In geography, the contemporary nature of the curriculum was reported to attract students, and the unitised approach to assessment to motivate students – as they can show their understanding and move on. One centre stated, however, that students appeared to find physical geography more interesting than human geography – and there was less of the former in the new specifications.

In mathematics, the unitisation of the specification was thought to have added focus for some students and to have motivated them, but pressure to perform in terms of league tables has meant that there was little time to engage students in what they might find inspiring and engaging.

For GCSE history, it was reported that the reduction in content has meant that there is more time for active learning approaches – such as role play, group work and discussions – which engage students.

Students’ own views about their level of engagement with, and enjoyment of, the different subjects were also mixed, although predominantly positive. Student comments suggested that they felt that the subject content of French, Spanish, geography and history was up to date, because the subjects helped them understand the world around them. Students

³⁷ Controlled assessment is a new form of internal assessment of the work of a course, replacing coursework.

reported having issues with the controlled assessment in French and Spanish, which are outlined later in this chapter.

4.1.4 Preparation for progression

For all six GCSE subjects, one or two centres reported that they had started to introduce new elements (including the development of subject-specific skills) into KS3 in the relevant subject, in order to prepare students fully for a GCSE course in the subject.

The introduction of controlled assessment in GCSE history was seen to prepare students well for A level but also to lead to fewer students being interested in taking the subject at GCSE. Two teachers from different centres were concerned that, although the current GCSE prepares students for A level, this is useful for only 20–30% of their students.

Several centres expressed the view that the new GCSE specifications for English, French and Spanish do not prepare students well for progression to A level, higher education or employment, because there was less emphasis on both subject-specific and generic higher-level skills development. However, there were some different views from the online survey which suggested that other centres believed that English and English language are likely to prepare students better for A level study. One case-study centre stated specifically that the international GCSE (iGCSE) in French and Spanish prepares students better for A level because it gives them more opportunities to develop as linguists.

4.1.5 Attainment

Much of the dialogue on attainment at GCSE was based on teachers' predictions for controlled assessment and examinations, as opposed to actual results. The overall feel was that attainment would remain at a similar level as before.

Where centres had had results for controlled assessments, these varied across and within subjects, although several centres reported a downward trend on their predicted grades for GCSE geography. One centre also reported surprising results for GCSE French controlled assessment, citing a high proportion of A*s and As, a few Bs and Cs and more Ds than expected.

In the online centre survey, only 22 of the 36 centres to report that controlled assessments had taken place (out of the 42 respondents) had received results for controlled assessment. Of these, 12 reported that the results were as expected, 4 that they were below expectations, a further 4 that they were significantly below expectations, and 2 that the results were above expectations. The subjects for which results were reported as being significantly below expectations were GCSE French and Spanish. The response size here is so small, however, that no reliable conclusions should be drawn at this stage.

4.1.6 Impact of controlled assessment on teaching and learning

The purpose of controlled assessment is to assess those aspects of a subject that cannot be easily assessed by external examination. These aspects include research, planning, investigation, analysis, collaborative working, and presenting ideas and arguments supported by evidence.³⁸ Controlled assessment is designed to encourage a more integrated approach to teaching, learning and assessment, and to enable teachers to confirm that students carry out the work involved.³⁹

Staff from seven case-study centres spoke about controlled assessment and how far they considered it to be integrated into their teaching and learning. The majority of controlled assessments take place within the normal timetabled lessons. However, there were four examples where students had taken or would be taking controlled assessments under examination conditions. Staff in general seemed to feel that controlled assessment was ‘taking more and more time from the course’, that ‘we’ve lost five weeks of teaching due to controlled assessments’ and that it also demanded a lot of time to prepare: ‘You spend your whole life preparing for controlled assessment,’ said one head of modern foreign languages. However, many modern foreign language teachers reported undertaking multiple controlled assessments to give students an opportunity to improve grades – only submitting the best grades to the awarding organisation. A member of the senior leadership team at one school – who appeared not to consider it learning when students worked independently while teachers gave assignments to others – was concerned that, when controlled assessments take place, ‘children are not getting actively taught’. Controlled assessment poses a particular challenge for French and Spanish teachers, because the speaking assessment has to be done on an individual basis. This, they feel, ‘detracts from the teaching and the enjoyment the children get out of it’.

A minority of teachers reported that they were integrating the controlled assessments into their teaching. One Spanish teacher considered themselves to be ‘flexible’, planning the controlled assessment when ‘something of interest crops up’ and they were ‘inspired’. A teacher of French at the same centre had recognised that they had got ‘too orientated towards controlled assessment’ and had decided to teach ‘normal skills’, as well as controlled assessment. A head of modern foreign languages reported that students ‘hate the assessment regime, because they can’t work out what is assessment and what isn’t’ – which the head saw as contrary to the ethos of controlled assessment. Another head of modern foreign languages and a headteacher had a contrary experience, in that they had found that

³⁸ QCDA (2010) Managing GCSE controlled assessment – a centre-wide approach.

³⁹ QCA (nd) Changes to GCSEs including controlled assessment, information for teachers (Ref: QCA/09/4174).

controlled assessment focuses students' learning, that students take it very seriously and that the quality of their work had improved.

Teachers select from a choice of topics and contextualise them for their own students, although the amount of choice varies between subjects. Most teachers were positive about the facility to choose and develop topics for their students. One teacher of English felt that controlled assessment fitted in well with their previous regime of 'assess(ing) every six weeks anyway'. A language department looked at controlled assessment over the two-year period, applying a unified approach to the three languages offered at GCSE. Another head of modern foreign languages felt that the ability to choose topics was a plus.

Controlled assessment replaced coursework in part because of concerns about the provenance of students' work. Some teachers continue to express concerns about how much help some students may be getting, particularly at home. One head of English was concerned that those students who could get help at home would have an advantage over those who could not. They also felt, with others, that their role in controlled assessment was a difficult one in that the feedback they can offer is limited and, again, disadvantages some students. One Spanish teacher felt that students could just regurgitate pre-learnt material. However, a head of French contradicted this by saying that, although students could learn certain things by heart, they still needed to answer questions that they could not prepare for in advance. One headteacher also reported that their staff believe that controlled assessment is better than coursework as 'you can be far more certain that it's genuinely the student'. Two centres were concerned at seeing students prepare work for use in the controlled assessment that was incorrect and not being allowed to help them. One geography teacher felt it was impossible to 'cheat' in coursework and that this aspect was therefore irrelevant to their subject. One respondent to the online centre survey had a different view, claiming that 'there is just as much cheating as before, if not more', and fearing that some teachers gave the tasks to their pupils for homework and then used exactly the same tasks for the controlled assessment.

4.1.6.1 Deep learning

A range of teachers was asked whether controlled assessments promote deep learning. There are concerns that controlled assessment may only test the ability to learn content and regurgitate it. In modern foreign languages, for instance, students are reportedly able to learn passages of their own writing and regurgitate them in the assessment room. One headteacher said that there was a concern that, in English, teachers can run a practice question and then just change the words in the title slightly and re-run it as the real assessment, thereby allowing students who memorise an answer to succeed. The same

point was made about history. The suspicion is that centres around the country will do this and that students will simply learn essays as part of the approach to controlled assessment. One teacher claimed that, in history, controlled assessment is all right for students with good memories but not for those who need to have their skills nurtured and developed with the assistance of their teachers – who are unable to intervene. A respondent to the Spanish online survey made a similar comment, that students ‘cannot remember enough’ to meet the word count required by the controlled assessment.

Other teachers also felt that this lack of allowed intervention does not promote real learning. Two centres, conversely, maintained that controlled assessments promote learning. One head of modern foreign languages stated that it makes students more focused in their learning. This view was echoed by a headteacher whose experience of controlled assessment is that it has focused their students’ learning and that students have coped well with it.

Interestingly, when students talked about controlled assessment in terms of a specific subject, they tended to be negative – but when they talked generally about it, they were very positive. Students of history and geography were the most positive about controlled assessment, and students of Spanish and French the most negative. The students suggested that modern foreign languages are not well suited for controlled assessment, whereas humanities (here geography and history) are. Geography and history students reported that they liked being able to take notes into the controlled assessment; they considered their skills were being assessed rather than their ability to just memorise facts. The majority of students also stated that they liked the fact that controlled assessment did not take place in the main school hall, which is where they sat external examinations, and that controlled assessment was therefore less stressful. Students of French, however, felt that controlled assessment was often just assessing their ability to memorise and regurgitate a text. Modern foreign language students also reported that the timescales for preparing for controlled assessment were much tighter than they were with other subjects.

4.1.6.2 Independent learning

Teachers disagree about whether controlled assessment promotes independent learning. One head of history feels that the underpinning assumption – that putting a group of students in a room, giving them a variety of sources and allowing them to talk, will produce independent learners – is flawed. In contrast to this, one head of geography and one head of English felt that some independent learning was being promoted and that the mark schemes also reward this skill. In French, one teacher described how the learning is now quite different, as students have to choose and present something that they are interested in. This

teacher's comments about students presenting their work confidently and competently seem to suggest a degree of independent learning – although this seems to be partially undermined in the second part of the assessment where, the same teacher says, students achieve less well, as they are given unexpected questions to answer spontaneously.

Students reported that they liked the opportunity to research and plan for controlled assessment and the fact that they were able to take notes in with them to the assessment, although students in one centre said that they did not know how to research, how much to research, how to write effective notes and how to use notes in the controlled assessment. Students from this centre also reported what they considered an unfair differentiation between subjects: in history, controlled assessments took place during lesson time so, if more research/notes were needed, this could be done between sessions; but geography took place over one day, so there were no further opportunities for research.

4.2 The impact of change on centre behaviour

4.2.1 Influence on centre choice of specification

4.2.1.1 Choice of specification

Many of the same reasons that were advanced for the centres' choice of specification at A level applied at GCSE, too. Two that were specific to GCSEs, however, related to the tiered structure of some GCSE courses – and, in the case of mathematics, to the advantages of a unitised rather than a linear model. In modern foreign languages, one centre welcomed the possibility they had of deciding – on the day – whether students did higher, or foundation, listening and reading. In mathematics, one centre entered students for the foundation-tier specification offered by one awarding organisation but for the higher tier with another awarding organisation. Also in mathematics, two of the centres were taking advantage of the switch from a linear to a unitised specification to increase – as they saw it – the chances of students' achieving higher grades.

4.2.1.2 Timing and frequency of assessments

The main issues raised in relation to the timing and frequency of assessments were:

- A concern that students were being over-assessed. The responses where this was raised related specifically to English and history, but there were indications that the concern was more widespread.
- Given the timing, one centre commented that it is difficult to make decisions about whether or not to enter a student for a re-sit, as the previous results are not known before new examination entries have to be made. Another centre noted,

with reference to geography, that not all awarding organisations offered January examinations and that this therefore limited the possibility of students' re-sitting.

- Referring to work in year 10, specifically mathematics, one centre felt that parents and students did not always appear to be aware that the examinations were summative. Students were therefore sometimes absent for an examination for unacceptable reasons – for example, because they were away on holiday.
- Year 10 examination entry – there was a concern about striking an appropriate balance between the maturity of the student and the point at which they are entered for an examination (in this case in history), so that the number of examinations they face is properly spread out.

4.2.1.3 Patterns of candidate entry

As with A levels, a majority of respondents felt that it was too early to assess whether the changes had had a discernible impact on candidate entry. With that caveat:

- three of the centres reported that there had been no discernible change
- in one centre the numbers taking modern foreign languages had increased, and the centre speculated whether this might be a result of the introduction of the English Baccalaureate – although, to counter this, another centre reported a decline in the numbers opting for Spanish
- in relation to history, one centre felt that it was generally perceived as being academically more challenging than geography and that numbers suffered as a consequence – another centre, however, reported that history was maintaining its popularity.

4.2.2 Management and resource implications for centres

In the online centre survey, there was evidence that changes to assessment had resulted in the development of new policies, particularly for controlled assessment. However, there was a difference in perspective between the senior leaders and the subject teachers on whether the new policies were at department level or at centre level. The view of the subject teachers is that policy changes have been primarily at the department level, whereas senior leaders believed that policy changes have been primarily at the centre level. It is possible that this indicates a certain lack of awareness of centre policy at departmental level, and of departmental policy at centre level.

4.2.2.1 Approaches to controlled assessment

The biggest impact on the GCSE centres has undoubtedly been the introduction of controlled assessments and the increased burden this has placed on management and resources.

QCDA guidance on controlled assessment states:

Every school and college must ensure that controlled assessment is managed effectively and operates smoothly to benefit students and comply with the regulations. Schools and colleges should introduce a centre-wide approach to controlled assessment from the start of teaching the revised GCSEs.⁴⁰

Of the 15 centres that completed some or all of the questions on controlled assessment in the centre survey senior leaders' questionnaire, most (13) centres reported having a centre-wide policy on controlled assessment. Of the 12 case-study centres offering GCSEs, 5 centres reported on this matter and, of those 5 centres, 2 have a centre-wide policy in place. Thus, in a sample of 20 centres, 15 centres have introduced a centre-wide policy or approach to controlled assessment.

A small number of the centres (in the case studies and survey) had a departmental policy. Hence, only 1 of 20 centres did not report having any policy in place to manage controlled assessment.

4.2.2.2 Managing the scheduling of controlled assessments across the centre

In the first round of centre visits in summer 2010, the majority of case-study centres liked the fact that they could choose when to hold the controlled assessments, although many were concerned about the implications of this flexibility on resources. In the second round of case-study visits in spring 2011, it emerged that arranging rooms, ICT and other resources were minor issues now, and only for a minority of centres. It appeared also that the initial difficulties experienced by a minority of centres, as reported in summer 2010, in catering for SEN students in terms of timetabling, staffing and rooming have been more or less resolved. Instead, a majority of centres reported that the most challenging aspect of the introduction of controlled assessment, from a management perspective, had been to accommodate all the required controlled assessments across the centre within an already tight school calendar.

Co-ordinating the timing of controlled assessments in different subjects so that they do not conflict poses a major logistical problem for the senior leadership team and departmental heads in schools. GCSE subjects might require two whole days for a controlled assessment,

⁴⁰ QCDA (2010): Managing GCSE controlled assessment. A centre-wide approach.

which in turn affects both the teaching time and available time slots for controlled assessment in other subjects. In a minority of centres, intense negotiations had been taking place between different heads of department seeking to secure a preferred time slot for their controlled assessment.

About half of the centres that had previously found that the scheduling of controlled assessments involved a significant added management burden reported now that they were coping well or satisfactorily with the issue. The remaining half thought that the challenge of fitting controlled assessments into the school calendar was becoming a more serious issue for them, as the number of departments involved was increasing.

Although departments were trying to timetable controlled assessments in the normal lesson time for each subject, in some cases this was not possible. It was reported by several teachers, for instance, that students were being taken out of other subjects' lesson time to do speaking assessments in French and Spanish.

One headteacher captured the sentiment of a number of staff members across different centres:

It is having impact on the whole curriculum because, if you are following the controlled assessments to the letter, then you are actually taking students out of other curriculum time in order to service the needs of a particular subject. [...] It seems we could end up with our noses buried in controlled assessment for the greater part of the year.

In the online centre survey, respondents were asked whether they felt that the balance of controlled assessment and external examination is good for their GCSE subject. Of the 42 respondents, 17 felt that the balance was good, and 17 that there was too much controlled assessment; 8 felt that they could not say. No respondent reported that there was too little controlled assessment. Looking at this result by subject shows that 15 of the 17 respondents who felt there was too much controlled assessment were commenting on modern foreign languages. However no general conclusions should be drawn due to the low response rate for the online centre survey.

4.2.2.3 Management of student absences

Centres and departments clearly need to arrange an alternative time for students to undertake a controlled assessment when they have, for a variety of reasons, been absent from the previously scheduled assessment. In summer 2010, 10 teachers reported that student absences during controlled assessments caused a great deal of rooming and staffing difficulties. In spring 2011, of the seven centres that discussed the management of

student absences, three found the issue problematic and four were coping with it reasonably well.

The first three centres thought that the management of absences added to the already existing challenge of scheduling controlled assessments. At these centres, students would have to miss lessons either in the subject in which they were taking a controlled assessment retrospectively, or in other subjects. The four centres that found the management of absences less problematic had opted for scheduling sittings out of lesson time, for example at the end of the day. This was seen as better for the students, but onerous for the teaching staff. Staff were doing more hours than they would do otherwise, often on a voluntary basis, and in some cases even giving up their lunch breaks to supervise controlled assessments. Providing cover for teachers who needed to be out of their normal lesson to supervise students taking their controlled assessment retrospectively was also mentioned as a slight management challenge.

Two centres reported that some students had an unhelpful attitude towards controlled assessment, which they had not shown towards exams. Students were more likely, for example, not to turn up for a controlled assessment on the grounds that they did not feel like it; or they would not tell their parents that they had an assessment on a specific day, and parents might schedule something (e.g. a medical appointment) that clashed with the time of the controlled assessment.

It was reported by a minority of teachers that managing disruptive students during controlled assessment was more difficult than during an exam, as the teacher was alone in the classroom and could not leave to find the disruptive student another room and a supervisor.

4.2.2.4 Staff development

As with A level, the majority of staff welcomed the opportunity to participate in training; and, again as with A level, most of the training available to them had been offered by awarding organisations. Training was especially welcomed when there had been changes to the specifications. One head of English wanted more support for staff on the spoken English element of the GCSE as well as on new areas of poetry. A head of mathematics was keen to provide more training for staff on functional mathematics. The same head of department said that they tended to send a key member of staff to training, who would then cascade what they had learnt to the rest of the department, as well as sharing their expertise among the team – they also shared this information with the schools in the cluster group in their area. Another centre, having identified problem solving as an area of uncertainty within mathematics, invited their awarding organisation to come to the school – ‘and that was really useful’. A head of modern foreign languages also said that training with clusters had been

useful. One form of training undertaken in a mathematics department had been collaborative planning. This approach had also been used at another centre in the geography department.

More especially, staff across the board – headteachers, heads of department and subject teachers – had found that they needed training to understand more fully what was expected of them and their students in relation to controlled assessment. They would have particularly welcomed clear advice on how to implement controlled assessment correctly (i.e. guidance on best practice in the management of controlled assessment). Teachers would also have welcomed further information on how final grade boundaries (controlled assessment/external assessment) were determined in different subjects.

Two teachers reported that awarding organisations had been providing new information and advice on controlled assessment when it was already underway – which had led to some confusion among staff. GCSE English students in one centre reported the same thing, saying that confusion had meant a lot of wasted class time.

5 Wider stakeholder perceptions of the impact of specification changes

The wider stakeholders (this group included subject associations, professional bodies and workforce development agencies) who were interviewed before last summer's A level results offered conflicting views, with some feeling that there is considerable comprehension and recognition of the impact the changes to specifications will have, and others that centres are not fully aware of or responsive to the changes. Some centres were thought to have developed a good understanding of the changes, particularly their focus on greater stretch and challenge. However, other stakeholders expressed concerns that centres perceived the changes to be minimal (the main change mentioned here was the move from six units to four). There had been fears that some centres had not picked up on stretch and challenge nor on the development of higher-order skills and independent learning.

As is to be expected from the wide range of organisations interviewed, responses reported were mixed, and the interviews focused on the particular area of interest for the stakeholder. Inevitably, views expressed by wider stakeholders are underpinned by individual perspectives on the 'what, how, why and for whom' of curriculum development.

The issues highlighted fell within two main categories: firstly, what constitutes the canon of recognised knowledge for a particular subject, the specialist language and the associated subject- or discipline-specific skills; secondly, why a subject should be taught, in the sense of whether it should prepare students for their next steps in life – into further study in an

academic or vocational context, or into work. There was some concern by subject associations about the imbalance of knowledge-based content and subject-specific skills. Other subject associations, however, welcomed the stretch in the new A level specifications and the synoptic skills required.

The evidence from the stakeholders enforces the academic/vocational split, with different skills and knowledge valued depending on whether these are considered work-oriented skills or whether their purpose is to develop an understanding of the subject for higher-level study. Employer priorities focused on employment. There was enthusiasm for the higher-level skills developed as part of the extended project qualification.

There had been a general expectation that the relatively steady pattern of attainment seen previously would be disrupted in 2010, partly because, at the time of interviewing, it was unclear what would be in the summer 2010 assessments for A2. It was considered too early to have any clear idea about attainment at GCSE.

In a follow-up meeting in June 2011 with one FE representative organisation, participants from colleges noted surprise that A level attainment in their colleges had not been affected more by the changes. Several mentioned that there had not been many significant changes to teaching and learning initially and that this had been expected to have a negative impact on results. Changes to support students to become more-independent learners were being implemented, including further development of resource centres where students could be supported with research skills. There was concern that, especially at A2, teaching had not been different enough to respond to the changes in the specifications. This reflected the views of the case-study centres that A2 required a significant jump in skills from AS level.

There was continued anxiety about the rate and volume of change teachers and learners are subjected to. Although much of the change to GCSEs and A levels was welcomed, it was seen as one of many changes implemented within a short space of time. This was felt to have an impact on centres' ability to plan effectively for the changes and to release staff for professional development. A silo effect was often reported, with departments or qualification teams planning in isolation, and with insufficient time and resources for more-effective whole-organisation approaches.

One awarding organisation expressed concern over the direction of travel for the relationship between awarding organisations and centres, with the awarding organisation providing ever-increasing support and the boundaries becoming blurred between the role of teacher training and independent provision of assessment.

Wider stakeholders saw the main driver for changes to GCSEs coming from the way in which coursework had become discredited. They had suggested that the main issue for the

future will be the outcome of the 'controlled assessment' approach, predicting that the lack of information about how controlled assessment is conducted, and the lack of support from awarding organisations on this matter, would have a negative impact.

Awarding organisations also considered that controlled assessment was a major workload for centres. They felt that, in many cases, centres' implementation is detracting from learning rather than providing an innovative and engaging experience for candidates that best enables them to demonstrate their capabilities in a subject. There was some belief, however, that controlled assessment can bring advantages when properly implemented and that a measured approach should be taken to revisions of the system, rather than any wholesale change, before the impact and best practice are fully understood.

The overall positive response to the introduction of stretch and challenge, with the benefits largely outweighing the small difficulties observed so far, was reiterated by the awarding organisations. However, unitised assessments with opportunities for retakes and early entry, coupled with pressure on centres and candidates to get results 'in the bag', were considered at odds with the value supposedly placed on synoptic assessment and on the development of skills rather than a narrow knowledge-based approach to subjects. There was concern that the overuse of early entry and re-sitting is detracting from candidates' overall learning experience, and is in turn placing more strain on the maintenance of a consistent examination standard targeted at a year 11 cohort.

The focus on assessment and assessment strategy is in danger of diverting time and attention from the high-quality delivery of teaching. Concerns were expressed that the process of five-yearly step changes in qualification design is unproductive and, indeed, damaging to teaching and learning.

The involvement of higher education in attributing value to the A levels was welcomed, and it was hoped that endorsements of quality from higher education may provide a counterpoint to pressure on centres to adopt 'easy' qualifications.

Of the 35 HEIs responding to the online survey, 20 said they had plans to change the admission criteria for the next academic year. Of these, 16 indicated that entry requirements were likely to rise for some or all courses, with around a quarter of these specifically mentioning competition for places as the driver; 6 pointed out that it is normal practice to review admissions criteria annually.

None of the HEIs responding said that their institution did not accept higher A level grades gained by students re-sitting AS or A2 units, but many listed certain caveats:

-
- some HEIs stated that, although re-sits are generally accepted, certain courses, notably medicine, would not accept re-sit results
 - some HEIs accept re-sits if these are sat within the two-year A level period, but students are discouraged or made higher offers if they re-sit after that period
 - some specified that re-sits may only be accepted if there are extenuating circumstances, although this is sometimes only a requirement for the more competitive courses.

Most HEIs said they consider a student's predicted or actual grade for the extended project qualification when making an offer. In the additional information respondents provided, many (approximately two-thirds) made it clear that they saw the extended project qualification more as a valuable piece of additional evidence than as a specific entry requirement. However, only a small number of HEIs (4) indicated that it would form part of an offer.

Of the 27 HEIs who provided an assessment of the accuracy of school/college predictions of A level grades in 2010, the majority thought that they were quite accurate. Of the 15 respondents who felt able to compare the 2010 predictions with the 2009 ones, most said there had been a similar degree of accuracy year on year.

The majority of HEIs indicated that more students applied for courses for entry in 2010 than in previous years. Almost half of the HEIs who responded in detail on this point attributed the increase in applications to the continuation of a recent trend for their own institution, often indicating that this was at least in part a result of their own efforts. Others indicated that they had seen a growth in admissions reflecting the national trend, with some specifically attributing the rise to the current economic climate and the upcoming rise in fees. A few HEIs indicated that they had seen a decrease in applications as a result of an increase in entry requirements.

6 Summary of preliminary findings

6.1 The impact of changes to the specifications on teaching and learning

Just over half of the wider stakeholders interviewed expressed concern that centres had not recognised the degree of change in the A level specifications – and some had particular concerns about the greater stretch and challenge required at A2 – and the subsequent consequences for teaching and learning.

'Stretch and challenge' is being interpreted in a variety of ways, and there is a degree of uncertainty and ambiguity about what is involved. Overall, however, the majority of centres felt that there was greater stretch and challenge in the new specifications at A2, although

this was not always equally reflected in the assessments. Physics was the only subject for which centres felt that there was less stretch and challenge than in the previous specifications.

It was reported that, to develop higher-level skills and independent learning in students, teachers need time to adjust their teaching strategies and students need time to develop new approaches in their learning.

The decrease in the number of units at A level from six to four (a decrease that does not apply to physics) was seen as positive in terms of ensuring greater depth and breadth of study.

Progression from AS to A2 was considered by many case-study centres to be a larger jump than in the previous specifications. During the follow up visits in 2011, for example, for English literature, history and geography those who responded thought that the AS was preparing students less well for A2 than before

Teachers felt that A2 English literature, geography, psychology, history and French were more challenging, and considered them better preparation for higher education – for example, in terms of encouraging independent learning. There were some exceptions in terms of specific skills development for a subject – for example, some geography teachers felt that there was not enough emphasis on the research skills required for university study.

Although the updating of course content had increased students' sense of motivation and engagement with GCSE subjects, this positive effect has been counteracted by the increased focus in teaching and learning on assessment.

6.2 The impact of the assessment regime on student development

A number of centres commented on the rise in numbers of A level students re-sitting examinations in order to achieve a higher grade. Students generally welcomed the opportunity for re-sits, with the majority believing re-sits should be allowed in any circumstance.

Although coursework was considered a burden for teachers in terms of their workload, it was also felt by many teachers to offer assessment which supported development (i.e. with a formative and summative element). Teachers, managers and students suggested that A level coursework performance was a better indicator of a student's ability than written examinations. Physics apart, the majority of A level students' subject- and non-subject-specific comments suggested that they would wish to retain coursework as part of the assessment regime.

There were some concerns about unitised assessment for GCSE. Five case-study centres discussed the respective merits of unitised and linear assessment in GCSE French, history and mathematics. They were of the view that linear assessment promoted more in-depth and longer-lasting learning than unitised assessment, particularly in relation to the development of subject-specific skills. Similar concerns were expressed by wider stakeholders. Unitised assessment with opportunities for retakes and early entry, coupled with pressure on centres and students to get results 'under their belt', was considered at odds with the value placed on synoptic assessment and on the development of skills rather than a narrow knowledge-based approach to subjects. However, this was balanced to some extent by teachers and students who felt there were benefits in allowing students to achieve along the way and to build on results.

The overall consensus from centres and awarding organisations was that there is conflict between the need to gain the 'results' required for school performance targets – using re-sits and early entry to maximise pass rates (e.g. at A*–C at GCSE), which is possible with unitised qualifications – and encouraging learning, development, coherence and a greater understanding of the subject.

Some case-study teachers reported that students are very outcome-focused and often reluctant to engage with any content that they do not think will be in the assessments. Stakeholders felt that the focus on assessment and assessment strategy diverted time and attention from high-quality teaching.

A wide range of approaches to controlled assessment⁴¹ is being implemented within and across centres, often depending on the subject being assessed and/or on the teacher. Awarding organisations have often interpreted Ofqual's generic guidance differently, meaning that there is sometimes variation in the advice provided for the same subject across the different awarding organisations. This has resulted in considerable variation in the amount of support that teachers believe they can give their students in the preparation stage of controlled assessment.

Concern was expressed that some teachers are using strategies, even if often unwillingly, to help students to get the best grade, and that these undermine the validity of the controlled assessments in terms of the knowledge and skills that are meant to be assessed. Examples of strategies include running practice assessments that change little in the 'real' assessments and modern foreign language teachers setting more controlled assessments than required and choosing the best for each individual student.

⁴¹ Controlled assessment is a new form of internal assessment of the work of a course, replacing coursework.

About half of the centres that had previously found that the scheduling of controlled assessments involved a significant added management burden reported now that they were coping well or satisfactorily with the issue. The remaining half thought that the challenge of fitting controlled assessments into the school calendar was becoming a more serious issue for them, as the number of departments involved was increasing. The majority of controlled assessment is undertaken during subject lesson time in many centres, although for modern foreign languages the need for assessment on a one-to-one basis means that students are sometimes taken out of other subject lessons. Many of the centres reported problems when there is a need to rearrange a controlled assessment for students who have been absent from the scheduled assessment, or when a student's behaviour means that they have to be asked to leave the room and miss part of the assessment.

It is accepted that change, such as the introduction of controlled assessment, initially increases teachers' workload and that this effect is usually only short term. However, teachers of modern foreign languages reported that they expect their workloads to remain heavier in the longer term also, because of specific requirements to do with organising and conducting controlled assessment.

Most teachers liked the opportunity that controlled assessment gives to choose topics and contextualise the tasks for their students. However, there was no consensus across or within centres and subject areas about the extent to which controlled assessment is an appropriate form of assessment for particular skills and knowledge. Controlled assessment was seen as summative assessment, with only a minority of teachers seeing it as having a formative role and being an integral part of the teaching and learning process. There are concerns that, rather than promoting in-depth independent learning, controlled assessment may only test the ability to learn content and regurgitate it.

When students talked about controlled assessment in terms of a specific subject, they tended to be negative – but when they talked generally about it, they were very positive. Students of history and geography were the most positive about controlled assessment, and students of Spanish and French the most negative. History and geography students liked the fact they could use their research notes during the controlled assessment. The majority of students also stated that they liked the fact that controlled assessment did not take place in the main school hall, which is where they sat external examinations, and that controlled assessment was therefore less stressful. There was some belief that controlled assessment can bring advantages when properly implemented and that a measured approach should be taken to revisions of the system. Staff at all levels expressed the need for guidance on best practice in the management of controlled assessment. Although awarding organisations got some credit for providing information and advice, a few teachers regretted that new

information was provided only when controlled assessment was already underway, which led to some confusion.

6.3 The impact of centre approaches on students

There is evidence from many of the case-study centres to suggest that the focus on school performance measures is a strong driver for centres to ensure students meet target grades. In many cases this means that teaching and learning is driven by the assessment regime, with centres relying heavily on awarding organisations for training to support students to pass examinations. There is evidence to suggest that some teachers and students will not consider specific content and/or skills that are not expected to be assessed as a priority, but rather as an indulgence. In some centres, however, the focus is less narrow, with centres embracing the change and challenge presented.

Choice of specification was reportedly determined by several factors, including the appropriateness for the needs and aspirations of students, consistency with the nature of the subject and familiarity with the awarding organisation.

6.4 Attainment in the context of the new specifications

The A* grade at A level is perceived as recognising high attainment by teachers and students. Some universities are starting to ask for an A* pass for certain courses.

Centres reported the use of a range of models for the management of the extended project qualification. The qualification was felt to have benefits for both student and staff development. Wider stakeholders expressed enthusiasm for the higher-level skills developed as part of the extended project qualification

The plateau in grades at A level seen in the statistical data⁴² does not fully reflect the increased level of challenge perceived by students and teachers before results from the summer 2010 examinations, especially at A2, where many teachers had expected to see a drop in attainment after the change of specification. The plateau in grades is likely to be due

⁴² From 1996 to 2009 (the last year in which the previous GCE specifications were awarded), A level attainment continued broadly on the same improving trend in most subjects – a small steady increase of about 0.06 of a grade per year in grades attained, punctuated by short-term dips downwards from the trend at times of specification change. Also, since around 2003, participation has been increasing in most subjects.

For results in 2010, a year in which another specification change has occurred (the specification change for teaching from 2008 working through to results in 2010), attainment saw a plateauing of grades following a long period of increasing grades – i.e. another shift downward from the trend, reminiscent of results in previous change periods in curricula (although much smaller in size than the dip in 2000/01). Looking at individual grades attained (across all new specification subjects combined), the same proportion of students obtained the A grade (including those who received the newly introduced A* grade) in 2010 as received an A grade in 2009 (24.6%), whereas, for example, from 2003 to 2009, on average year on year 0.7% more students received an A grade.

to measures introduced by Ofqual to ensure that the first candidates to complete the new specification A levels did not suffer disadvantage as a result.

Analysis of demographics, prior attainment and the impact of re-sitting in the statistical data does not offer an explanation for the overall plateau in A level grades for the new specifications that occurred in summer 2010. Changes in the type of candidate taking the new specification A levels may provide the most obvious explanation for a change between that year and previous years in terms of grades attained. The plateauing of grades from 2009 to 2010, however, cannot be explained by factors relating to the candidates' background. Historically A level grades have been increasing year on year, with a similar trend of increase to the grades achieved by candidates at GCSE; and yet, if the GCSE results of 2010 A level candidates are compared with those taking A levels in 2009 and 2008, the 2010 candidates' GCSE results are similarly higher, but the A level grades they attained have not increased. In addition, the demographic make-up of the 2010 A level candidature is not demonstrably different from that of the 2009 candidature.

Concerns about an expected drop in attainment suggest that centres may not fully understand how the awarding principles of maintaining standards are intended to support a gradual change; they also suggest that awarding organisations are required to manage change in ways that do not disadvantage a particular cohort of students (i.e. those who are facing the changes for the first time).⁴³

6.5 Determining factors for driving educational change

The impact and speed of change seen to date is determined in part by differences in:

- the stage of implementation – i.e. whether centres have already completed a full cycle of the new specification
- the extent to which centres recognised change in the specifications and in the method of enquiry these encourage
- where teachers and students were in terms of adjusting to new approaches to teaching, learning and development
- how centres and/or individual teachers perceived the role of assessment in teaching and learning and, in particular, the extent to which teachers felt the assessments were compatible with, and valid for, the subject and the skills to be assessed

⁴³ The prime objectives of maintaining grade standards over time and across different specifications within a qualification type (Ofqual, April 2010, *GCSE, GCE, Principal Learning and Project Code of Practice*, para 6.2) necessarily become more problematic, and concerns among stakeholders more likely to arise, at times of curricular change.

-
- how pressured centres and teachers felt to ensure that, overall, their students gained the desired pass rates and grades to meet national targets
 - expectations and perceptions of achievement for particular student cohorts and/or profiles
 - how confident centres felt in managing the change process, their capacity, and the support offered and taken up.

It is possible that these differences would be the same for any curriculum and/or specification change and, overall, would be expected in these early stages of implementation. However, there is strong evidence to suggest that new specifications alone cannot drive educational reform. A range of other drivers and context related variables such as the focus on assessment and performance targets are impacting on centre behaviour in terms of teaching, learning and development, and 'skew' classroom practice.

Explanations for the plateau in A level grades awarded in 2010 can only be tentative at this stage. As teachers and students adjust to new approaches in teaching and learning, further statistical data analysis is needed, following the summer 2011 assessment results, to see if this plateau is a result of the 'newness' of the specifications. Additional data from centres will allow a more detailed exploration and understanding of the results in terms of specific cohort groups, including students who need longer to develop independent learning skills or greater support with conceptualised approaches.

Although there is some suggestion from centres that the new specification at AS is discouraging progression to the full A level in specific subjects at a local level, further analysis of the statistical data is needed to see the extent to which this is reflected nationally.

A more complete picture of attainment for GCSE will be possible following analysis of the summer 2011 statistical data. The high level of interest in the impact of unitisation and controlled assessment on teaching and learning has made this a key focus of this report. Fieldwork in the autumn term 2011 will offer an opportunity to look in more detail at the impact of other changes to the specifications.

Ref: DFE-RR170

ISBN: 978-1-78105-027-9

© AlphaPlus consultancy Ltd

November 2011