

Price £15

The development of e-assessment  
2004 to 2014  
by

The Exam on Demand Assessment Advisory Group

THE SECOND OF A SERIES OF OCCASIONAL PAPERS  
ON E-ASSESSMENT ISSUES

sponsored by

Exam on Demand Ltd  
September 2005



**EXAM ON DEMAND**

Exam on Demand Ltd  
Broad Farm Business Park, North Street, Hellingly  
East Sussex BN27 4DU

Series Editor: Dr Michael Kingdon  
Chair: ExoD Assessment Advisory Group

# The development of e-assessment 2004 to 2014

by

The Exam on Demand Assessment Advisory Group

THE SECOND OF A SERIES OF OCCASIONAL PAPERS  
ON  
E-ASSESSMENT ISSUES

sponsored by

Exam on Demand Ltd

## **Exam on Demand and the Assessment Advisory Group**

---

*Exam on Demand Ltd (ExoD) is one of the leading providers of on-screen, on-demand qualifications delivery systems to awarding bodies. The company is also the principal provider of such systems for use in prisons and other secure institutions.*

*As a major provider of e-assessment systems to UK awarding bodies and other agencies, ExoD has established an Assessment Advisory Group. One of the ongoing functions of the Group is to produce research and discussion papers to foster professional and public understanding of e-assessment issues. The Group is chaired by Dr Michael Kingdon and draws its membership from staff in schools and further/ higher education, plus consultants with experience of assessment at the awarding body and regulator levels.*

*This second paper takes the form of a hypothetical report produced in 2014 that looks back over the ten years since the publication of QCA's April 2004 blueprint and timetable for the implementation of e-assessment. Its focus is the personalisation of learning and assessment.*

## Contents

<b>Introduction</b> .....	3
<b>Part A:</b> QCA and the regulation of qualifications, 1980 to 2004 .....	4
The situation in 2004.....	7
<b>Part B:</b> Key e-assessment developments, 2004 to 2014 .....	9
E-testing and e-marking of vocational qualifications .....	10
E-testing and e-marking of general qualifications.....	11
The development of on-screen, on-demand AS/A-level examinations .....	13
The wide scale adoption of e-assessment.....	13
The availability of e-learning materials .....	13
Developing the resources for e-assessment .....	15
<b>Part C:</b> Features of the assessment system in 2014.....	17
Registration for GCSE and AS/A-level qualifications .....	17
Syllabus delivery media .....	17
Question types and associated marking methods .....	18
Banking of questions, section and papers.....	18
The hierarchy of assessments .....	19
Supporting and monitoring the delivery of e-assessments .....	20
Awarding of on-demand qualifications.....	21
<b>Part D:</b> The case studies .....	22
<b>Part E:</b> Outcomes, implications and projections.....	27
Changes to the nature of the GCSE and AS/A-level examinations .....	27
QCA and the regulation of e-assessment .....	27
Implications for stakeholders .....	28

## Introduction

---

In April 2004 Dr Ken Boston, the Chief Executive of the Qualifications and Curriculum Authority for England (QCA), published a five-year programme for the implementation of on-demand e-assessment. Key targets<sup>i</sup> included:

- 75% of the Basic and Key Skills tests for Levels 1 and 2 of the contemporary national qualifications framework (NQF) were to be delivered on-screen by 2005;
- Field trials of the first on-screen GCSE subjects were to begin in 2005;
- The three English unitary awarding bodies were expected to offer the first on-screen GCSE examinations by 2006;
- Codes of practice plus audit and regulatory criteria were to be in place for 2007;
- 10% of GCSE examinations were to be delivered on screen by 2007;
- Introduction of the first on-demand GCSE examinations was timed for 2008;
- On-screen, on-demand delivery of GCSE examinations was to be the norm by 2009;
- (Although a separate timetable was not provided, it was implicit that on-screen, on-demand assessment of units of AS/A-level<sup>ii</sup> examinations and many other qualifications would follow.)

Since his appointment in 2002, Boston had been describing the marking of the national curriculum key stage 2 and 3 (KS2/3) tests, the GCSE and AS/A-level examinations as a "cottage industry". He reasoned that new ways were required to focus markers' work, reduce their clerical burden and reform the time scales to which they worked. E-assessment, especially the on-screen, on-demand delivery of qualification units, was seen as the means of reforming contemporary assessments, examinations and qualifications system into something that better met the needs of learners and other stakeholders.

QCA accepted from the start that, eventually, use would be made of automatic rating of students' extended written answers, with on-screen marking by clerical or expert markers when automatic rating systems were not considered to be appropriate. Further, awarding bodies would accept and assess e-portfolios from an early date. To support these changes, it was acknowledged that many aspects of the GCSE, and subsequently other qualifications, had to change.

This second occasional paper in the Exam on Demand series takes the form of a hypothetical report, produced in 2014, which explores:

1. Why so few of QCA's 2004 e-assessment targets were achieved, including how over-regulation during the period 1980 to 2004 had stifled innovation in assessment and hampered the implementation of e-assessment;
2. Why the eventual success of the e-assessment was due mainly to the concomitant development of e-learning;
3. The factors that continue to shape e-assessment in 2014.

The focus of the paper is the e-assessment of general qualifications - the GCSE and AS/A-levels. Developments in vocational qualifications are reported, where relevant. Case studies are provided of how three secondary schools and a further education college responded to the challenges of e-assessment and e-learning.

## Part A: QCA and the regulation of qualifications, 1980 to 2004

---

Boston had anticipated that the regulator's relationship with the awarding bodies, especially the three unitary ones, would have to change if e-assessment was to be introduced to the timetable he outlined in April 2004. However, he and the later chief executives must have been surprised just how much of the regulatory structure that had been put in place since 1980 had to be unwound.

Prior to the mid 1960s, the typical GCE O-level examination for most subjects consisted of a single written examination paper, with practical work for some of the sciences and orals for modern foreign languages. During the next 15 years not only did the number of examined subjects increase, the number and types of techniques used to assess individual subjects increased, and the number of units per subject increased to 2 or 3 in response to:

- The Nuffield, Schools Council and other syllabus development projects;
- The use of multiple-choice questions (MCQ) and other new assessment techniques. Although MCQs were introduced as part of the Nuffield science teaching projects, they were used across the curriculum including some language comprehension testing for English and French O-levels;
- The extensive use of teacher assessment (TA) by the CSE boards.

Because of the strong links between A-levels and university matriculation, faculty staff in the universities that managed the GCE examination boards made significant contributions to the development of the overall systems and individual subjects, often giving their time for free. From the mid 1960s the GCE boards appointed research teams to manage new initiatives such as multiple-choice testing, conduct inter-GCE board work on subject comparability and enable the boards to develop collective views on contemporary assessment issues.

By the late 1970s an English secondary school was able to exercise considerable choice over what was to be taught for each subject and how it was to be examined. Depending upon the subject, a secondary school could expect to choose from the following syllabuses for its 16+ students:

- One or more Mode 1 syllabuses on offer from whichever of the 13 regional CSE boards was local to the school;
- (Possibly) CSE local area Mode 3 syllabuses;
- School developed Mode 3 syllabuses;
- One or more GCE O-level syllabuses offered by each of the six English school examination boards;
- (Depending on the subject) O-level syllabuses based on Nuffield, Schools Council and other development projects, examined by one GCE board on behalf the others;
- Even, early Joint 16+ syllabuses, forerunners of the GCSE, organised by GCE O-level and CSE examination boards working together.

Choice at 18+ was restricted to the GCE A-level syllabuses offered by the six English school examination boards and those based on Nuffield, Schools Council and other development projects. Even this degree of choice was perceived to conflict with the needs of higher education and employers. In response, the GCE examination boards standardised the content of large entry A-level subjects by the creation in 1983 of "common cores" that formed the major proportions of the subject content and examinations. At about the same time the GCE A-level awarding bodies produced good practice guides, which they all pledged to observe. For the CSE examinations<sup>iii</sup>, all of these initiatives tended to be centralised under the Schools Council.

During the 1980s, school, and later vocational, qualifications in England became ever more tightly regulated. Under the GCSE General and Subject-Specific National Criteria for GCSE Examinations,

published in 1985, the number of subjects to be examined was reduced, the content and approved assessment techniques for the allowed subjects became tightly constrained, and the number of organisations authorised to deliver the new examinations in England was reduced to four regional groups. A decade later these groups were reduced further to form just three English "unitary awarding bodies", authorised to provide GCSE and AS/A-level examinations.

The development of the Advanced Supplementary (AS) level in 1989 initiated a similar process at 18+. As part of this process, a forerunner of QCA<sup>iv</sup> took over the definition of the AS/A level common cores and the codes of practice from 1986. The regulation of vocational qualifications began the same year. A single method of awarding A-levels was imposed in 1987 and QCA absorbed the GCE boards' codes of practice and made them mandatory. Subsequent revisions of the GCSE and AS/A-levels during the 1990s further constrained the range of allowed assessment techniques and content, progressively reducing the differences in the syllabuses offered by individual GCSE and AS/A-level awarding bodies. Multiple-choice questions were banned in the early 1990s; teacher assessment was reduced in stages during the first half of the decade. A consequence of the new levels of regulation was that the involvement of university staff in the development of the GCSE and AS/A-levels declined, and the links between school and university courses were weakened.

There were attempts, before and after the introduction of the GCSE, to integrate graduated learning schemes with external assessments. Examples included:

- The Kent Mathematics Project, which developed graded study tasks for each of the main topics of the subject. Progress to the next task was dependent upon completion of end of task tests drawn from a bank. The tests were teacher assessed and moderated by LEA inspectors and project staff. The learning of the subject became personalised as students progressed through advanced, supporting and enrichment units at their own pace. GCE O-level and CSE awards were made by the University of London School Examinations Board and the South East Regional Examinations Board, respectively, on the basis of the depth and breadth of study achieved and a single-paper summative examination;
- The Graded Assessment (GA) movement. This developed learning/ assessment schemes in mathematics, science, modern languages, English and other subjects. The courses and assessments were developed by Kings College, University of London and the University of London Examinations and Assessment Council (ULEAC). Assessment was based on moderated teacher assessments and external tests, which led to GCSE awards from ULEAC.

The Kent Maths project was wound up when the GCSE was introduced because it did not comply with the General and Subject Specific National Criteria for the examination. Graded Assessments did not receive support from QCA's forerunners and were discontinued by the mid 1990s.

One area of development at GCE AS/A-level that was later to assist their e-assessment was the unification of syllabuses and assessment. This started as an initiative by a single school in the mid 1980s, proved immediately popular with others and was adopted by QCA as the standard model for AS/A-level syllabuses during the revisions that took place during the early 1990s. To demonstrate their understanding of the overall A-level subject content, QCA required all students to take a final (synoptic) assessment.

National Curriculum testing was introduced progressively from 1990 and was managed by QCA. The introduction of the key stages 2 and 3 (KS2/3) tests in 1995, impacted on the relationship between QCA and the unitary awarding bodies. The latter were allowed, with others, to tender for test development, marking and data collection services. It was perhaps inevitable that, towards the end of the 1990s, the idea began to circulate in regulatory and government circles that the costs of the GCSE

and AS/A-level examinations would be reduced and their efficiency increased if they were also managed centrally. The proposal to create an English national awarding body, based of course on QCA, was contained in the conservative and labour party manifestos for the 1996 general election.

Many QCA and awarding body staff expected the national body to be created sometime during the second Blair government. However, several very public failings in the UK examination system served to dissuade ministers from taking too direct a responsibility for school examinations:

- In 2000 the SQA (the Scottish version of QCA and the national awarding body for Scotland) failed to deliver a significant number of results following a revision of the Scottish "Highers" examination;
- In 2001 the first awards for new A-level examinations took place and there were accusations in the media that the awards had been pegged artificially to the 2000 levels. All three of the unitary awarding bodies claimed that the numbers of assessment units per subject had increased to a level at which they were experiencing difficulties appointing the required numbers of markers. Perhaps it is not surprising that, also in 2001, one of the three unitary awarding bodies failed to deliver a significant number of its AS/A-level results on time.

The consequence was that QCA remained as a buffer between ministers and the awarding bodies. Further, its first initiative to reform general qualifications management was markedly unsuccessful:

- In 2004 part of QCA became the National Assessment Agency (NAA). Its functions were to professionalise the work of markers, to address the issue of marker supply and be the awarding body for the National Curriculum assessments. In its very first year of operation the NAA failed to deliver many KS3 English test results on time.

A common pattern of failure, involving some or all of the following factors, became evident:

- A major overhaul of an examination or assessment that:
- Increased the number and radically changed the nature of the assessment units;
- Required the appointment of many new markers;
- Required new methods of awarding and/ or reporting results;
- The creation of a new national agency to manage the change process;
- An attempt to impose a managerial paradigm that conflicted with established assessment principles;
- The failure to trial new assessment features adequately.

In 2005, a year after QCA's blueprint for e-assessment development, New Zealand provided another example of a managed assessment failure. January of that year saw the publication of the first results for Level 4 (the former scholarship level) of the New Zealand Certificate of Educational Attainment (NCEA), success in which led to the award of university bursaries.

Awards for all four levels of the NCEA were based on a criterion referenced system. The first results for Level 4 proved to be neither valid nor reliable, evidenced by the absence of comparability between subjects and over time. The pass rates for languages differed by several orders of magnitude from those for the sciences. For example, 75% of the candidates that took Te Reo (Maori) passed, as did 74% for Chinese. The lowest language pass rate was 36.3% for Spanish. Pass rates for the sciences ranged from 9.1% for chemistry to 1.4 % for biology! Collectively, the new subject examinations proved to be more difficult than their predecessors and the overall pass rate at Level 4 dropped to half that of the previous three years. The short-term solution was the creation of one-off government bursaries for 2005 and the longer-term one was the return to a norm-referenced system for 2006 onwards.

## The situation in 2004

By the end of 2004 ten awarding bodies were offering on-demand and e-assessed versions of units for some of their lower level vocational qualifications. Further, QCA was supporting the on-demand e-assessment of the Basic and Key Skills qualifications. The tests being presented were predominantly multiple-choice in format, and had been written for delivery on paper. While this could be said to be a start, it was a limited one. The delivery systems being used were generally awarding body specific. Only one organisation<sup>v</sup> offered a common delivery platform to different awarding bodies.

The use of e-assessment by the universities and some chartered professional bodies was developing independently, and largely outside QCA's regulatory constraints. E-assessment systems were well established as components of e-learning systems for degree courses in medicine and professional courses in finance.

As Boston was proposing his blueprint and timetable for e-assessment, there were a number of regulatory issues that, together, provided an unsteady start to his e-assessment initiative. First, the homogenisation of the content and assessment from 1980 onwards meant that the GCSE and AS/A-level syllabuses of 2004 were to prove an inflexible basis from which to begin e-assessment development. The ongoing ban on the use of multiple-choice testing for these examinations was just one example.

Second, innovation in GCSE AS/A-level assessment and awarding techniques had effectively been centralised by QCA in the early 1990s and restricted soon after. When new assessment and award models were required for e-assessment, the most relevant experience was from pre-1990.

Third, there were the issues of awarding body finance and risk. The unitary awarding bodies had funded the successive rounds of GCSE and A/AS-level syllabus development. This meant that they did not have the reserves to fund further rapid change. Instead, the funding of e-assessment developments had to come from entry fees. As a consequence, developments were gradual, with each awarding body balancing the costs with the commercial risks of being either too early or too late into the e-assessment market.

Fourth, there were concerns that QCA might initiate a further round of GCSE syllabus developments concurrently with the on-screen and on-demand developments. Given the costs and distraction that this would have caused, the implementation of e-assessment could have been delayed for years. Fortunately other than implementing the recommendations of the Smith Committee for post-14 mathematics, QCA was persuaded to place a moratorium on other GCSE developments for the remainder of the five-year timetable. Smith et al (2004)<sup>vii</sup> had recommended a reduction in the number of tiers for GCSE mathematics from three to two, with the first examination in 2008.

The involvement of the universities in GCSE and AS/A-level courses development and examining had reached an all-time low. By 2004 some faculty admissions officers had come to believe that the value and relevance of these qualifications as tools for university admission had become so weak that new, university managed, entrance tests were needed.

Further, October 2004 saw the publication of the Tomlinson Report<sup>vii</sup>, which proposed a four-level system of qualifications for the 14 to 19 age range. The report advocated the reduction of external assessment in favour of institutional-based assessment and moderation, so was popular with teachers groups. Doubts were raised, by awarding bodies and others, about the reliability and comparability of the assessment methods proposed<sup>viii</sup>. The Report was also lukewarm to the concept of e-assessment and some of its recommendations conflicted with QCA's blueprint published earlier that year.

At first little attention was paid in the UK to the 2005 New Zealand problems. Later, it was pointed out ix that the NCEA shared common features with Tomlinson's recommendations and therefore raised questions about their viability. The NCEA failings illustrated the very concerns that ExoD (2005A) and others had expressed.

Although, the Government comprehensively dismissed the Tomlinson Report, the debate that followed served to reinforce several factors that were to prove critical to the eventual success of e-assessment:

1. The importance of the years 14-19 as a coherent stage of education and training;
2. The artificiality of the distinctions that had arisen between GCSE and AS/A-levels methods of teaching and assessment on one hand and between academic and vocational qualifications on the other;
3. The need to personalise the courses provided to students aged 14 to 19.

Also on the positive side, 2004 also saw the first pilots of the KS3 ICT tests and the first national pilot followed in 2005. To avoid the problems associated with the contemporary Internet, and to standardise the contexts in which KS3 students were assessed, the tests used a virtual net that had been created for the purpose. The project was a mammoth undertaking for QCA and its contractors, and suffered very few of the problems that bedevilled other government IT projects of the period.

## Part B: Key e-assessment developments, 2004 to 2014

---

The implementation of e-assessment 2004 to 2014 involved five separate applications of ICT and multimedia to assessment:

1. E-testing;
2. E-marking;
3. E-portfolios;
4. E-vivas;
5. E-presentations.

The term e-assessment was used from 2004 for the on-screen equivalents of traditional examination processes. For the purposes of this paper, it will be considered as two separate processes - e-testing and e-marking.

E-testing became the standard method for the delivery of assessment "papers" on-screen to students and electronic capture of their responses. The term was applied initially to external on-screen tests that contained multiple-choice and other short answer questions that could be scored automatically using optical mark reading (OMR) and intelligent character reading (ICR) technologies. As e-assessment and e-learning developed, the term came to be applied to all tests, assessments and examination papers that are presented on-screen, no matter what mix of questions they contain, whether they were internal or external, or how they are marked.

The term e-marking was initially applied to the marking of on-screen images of students' responses produced by electronically scanning their examination scripts. The process was already established by 2004 and overseas markers were employed as early as 2005. Later, imaging of examination scripts declined as greater quantities of students' responses were captured on-screen. As this change took place, the term e-marking became used to cover three different approaches to marking:

1. Automatic scoring of coded responses using OMR and ICR, and simple words or phrases using ICR;
2. Automatic rating of content and/or language structure;
3. On-screen marking by clerical and expert markers of extended written responses.

Even before 2004, expert markers were used to develop and audit the automatic (OMR/ICR) scoring of MCQ and other objective questions. Clerical markers were also being used to mark non-statutory key stage assessments. Systems for the automatic rating of content and language structure were in development in the US and elsewhere in the world. When a few years later these systems were used in the UK, expert markers were also used to develop and audit individual applications of auto-rating systems. Expert markers also became the standard against which double marking systems, using combinations of clerical and auto-rating pairings, were evaluated. All remarking for appeals and cases for special consideration were to remain reliant on expert markers.

E-portfolios rapidly became the standard means for the compilation, assessment and moderation of the reports, special studies, field and course work notes, retained samples of work, projects and records of competencies achieved that are required for current qualifications. By 2009 approved e-portfolio systems enabled students to incorporate a wide range of media, provided audit systems for teachers and assessors to monitor students' progress, and assisted with the authentication of the work submitted.

E-vivas had also been piloted by 2004. Even as early as 2008 they were being used to provide assessors and moderators with a means to question students directly, using mobile phones or video conferencing, and record the dialogue between them for e-rating and audit purposes. Similarly, e-

presentations developed to provide a means for assessors and moderators to observe, record and assess directly students' performances in the performing arts and practical units of other qualifications, without the need for the candidate or assessor to travel.

### **E-testing and e-marking of vocational qualifications**

Given the tightly regulated nature of the 2004 GCSE syllabuses, it is not surprising that early e-assessment developments took place in the vocational area. The vocational qualifications of 2004 also tended to contain fewer assessment units. Indeed, many of the lower level examples consisted of single units. The established systems of vocational assessment were also more amenable to early e-assessment. In particular, the ban on the use of multiple-choice questioning (MCQ) that QCA's predecessors had imposed on the GCSE and AS/A-level examinations, had not been applied to vocational qualifications. Therefore, the vocational awarding bodies already had banks of MCQ tests and items to support on-screen, on-demand assessment.

The vocational awarding bodies were also the first to develop e-learning materials. As early as 2005, one of the largest vocational awarding bodies was developing e-learning materials for each of its large entry own-brand qualifications. Professional bodies, training organisations, sector skills councils, even the DfES contributed to the development of e-learning materials for particular vocational qualifications. In contrast, any development of learning materials for GCSE and AS/A-level courses by the unitary awarding bodies was considered by many to conflict with their prime function of external assessment. This attitude did change but not until after 2008.

The fact that the vocational awarding bodies were answerable to other masters - other government ministries, the learning and skills councils, sector skills councils and professional bodies - as well as QCA, undoubtedly helped. Many of these other agencies were able to provide influence and resources to ensure that e-assessment progressed. The implementation of e-assessment can also be said to have been easier for vocational centres. They tended, given their size and roles in 2004, to have the necessary resources to implement e-assessment and the timetable flexibility to manage on-demand assessment.

Between 2006 and 2008 the vocational awarding bodies were seen as the greater innovators over the unitary awarding bodies. Some of those schools and colleges that were able to support e-assessment began to use vocational qualifications, simply because there were e-assessed.

The downside of these early advances by the vocational awarding bodies was the poor quality of the on-screen tests. Many had been written for on-paper delivery, with minimal adaptations for the new medium. The poor quality of these early on-screen vocational tests may also have served to reinforce some teachers' prejudices against the on-screen delivery of GCSE and AS/A-level examinations.

The vocational awarding bodies found plenty of opportunities to move their large-entry and lower level qualifications to e-testing, without the need to become involved in early work on on-screen e-marking and automatic rating of content. Until well after 2008, multiple-choice and short answer questions, that could be scored using OMR and ICR systems, predominated. Even when automatic rating of answers became a standard technique among the unitary awarding bodies, the smaller numbers of students that studied many vocational qualifications meant that the technique remained uneconomic. E-marking by clerical and expert markers was not widely used in the vocational domain until the higher level qualifications were assessed on-screen and on-demand.

## E-testing and e-marking of general qualifications

When the unitary awarding bodies began the development of on-screen versions of examination for GCSE subjects, the expectation was that there would be three broad stages to the process:

1. Adaptation of the examinations for existing (2004) syllabuses for delivery on-screen, on pre-published dates, with some students' responses being captured on-screen and some on-paper;
2. Adaptation of the examinations for existing (2004) syllabuses for delivery on-screen and on-demand, with some students' responses being captured on-screen and some on-paper; and for some subjects only,
3. Extension of the on-screen and on-demand examinations so that all responses could be captured on-screen.

Initial progress with the development of on-screen delivery was slow and the pattern that appeared was very different from that expected. The first task of the unitary awarding bodies was to identify those GCSE subjects that could be put on-screen without major restructuring of the 2004 syllabuses. Examples included the modern languages, geography and the sciences. The issue for the other subjects was that their existing assessment units contained sections that could not be assessed on-screen using contemporary e-assessment techniques. If the existing assessment structures had to be used, the only option was to sub-divide them further into the sections that could and could not be assessed on-screen and deliver them to students separately. After an extended consultation QCA conceded that, to avoid increasing the number of units to be assessed, the assessment structures for some subjects could be restructured to facilitate on-screen assessment. This enabled the development of on-screen GCSE examinations for history and other humanities subjects to begin, and the unitary awarding body that had achieved the least progress with its science syllabuses, took time out to restructure them as well.

As the range of question types that could be assessed on-screen developed after 2006, two patterns of e-assessment became the norm for GCSE subjects. First, for those (Pattern 1) subjects that did not have natural distinctions within the subject content, the schools preferred to see individual assessment units differentiated by question type and how they were to be e-marked, for example:

- Assessment units containing MCQ and short answers that could be scored automatically;
- Assessment units containing longer, often content rich, answers that could be marked using combinations of automatic rating (of content or language structure) and human (clerical) marking;
- Assessment units requiring extended writing to be marked by expert markers.

For the remaining (Pattern 2) GCSE subjects that do have natural distinctions within the subject content, as between - say - human and physical geography, the schools expressed strong preferences to have the individual assessment units differentiated to reflect these natural divisions within the subject content. Similarly, the unitary awarding bodies' preferences were then to have the individual P2 assessment units divided into sections that contained the different question types, and their associated e-marking techniques.

The awarding bodies began to develop question and paper banks, with section banks for the Pattern 2 subjects, from 2006. Also in 2006, e-assessed versions of several modern languages, the sciences and geography were in development and the first of these subjects were being piloted. Progress with mathematics - at first sight an obvious candidate for e-assessment - was hampered by QCA's decision to reduce the number of tiers to two and by issues of on-screen notation. As a result, the first on-screen pilots for mathematics did not take place until 2008.

The reactions of schools to e-assessment became increasingly important from 2007 onwards. For several reasons schools were reluctant to pilot the first on-screen GCSE examinations, especially the non-live ones. First, many lacked the resources or confidence to arrange large-scale, simultaneous delivery of scheduled on-screen GCSE assessments. Second, while individual delivery using a small number of workstations (sometimes referred to as "hot-desking" in QCA reports) was always an option, schools also remained reluctant to disrupt their timetables for non-live pilots. Finally, after the failures of many early on-screen sessions, schools demanded that all pilots, especially live ones, were adequately backed-up against technical and other failures by early opportunities to present an alternative test. This requirement was not new and replicated the policy that had been agreed for the Basic and Key Skills on-screen pilots conducted in 2003 to 2004x.

Also during 2006, it became clear that the wide scale development of e-assessment, for both academic and vocational qualifications, would be dependent upon coordination of at least five factors:

1. Agreement of the on-screen question types that could be used to assess individual GCSE and AS/A-level syllabuses, plus vocational qualifications;
2. Regular reviews of the regulatory structures to maintain relevance and flexibility;
3. Supply of good quality e-learning materials to support personalised learning;
4. Development of technical, physical and human resources by schools and colleges to manage e-assessment;
5. Schools and colleges developing the organisational structures to support personalised learning and on-demand assessment.

Once the first on-screen questions types were agreed with QCA, development on the question, section and paper banks to support on-demand e-assessment of GCSE subjects began. As part of its actions to assist the achievement of its timetable, QCA lifted the ban on the use of multiple-choice questions in general qualifications in 2008. Following from this decision the expectations of most stakeholders were that the banked questions would be pre-tested and that traditional post hoc awards would give way to pre-hoc calibration of section and papers, based on the psychometric properties of the questions they contained.

The list of agreed question types proved to be dynamic. New ones were constantly added and some of those on the list proved to be psychometrically inefficient, too expensive, or unpopular with teachers. As a result, e-syllabuses for the same subject, developed at different points in time, inevitably used differing combinations of question types, providing schools with choices about how individual subjects should be assessed.

From their responses to the questionnaire studies conducted by QCA and the awarding bodies as part of the first on-screen pilots, it was clear that many teachers were overtly hostile to the concept of e-assessment in general and on-screen, on-demand assessment in particular. However, as they came to accept the principles of personalised learning and appreciate the value of key features such as e-learning, when-ready assessment and the feedback that could be derived from e-assessment, the contribution that e-testing could make to personalised learning was gradually acknowledged.

Teachers' reactions also varied with their subject. While teachers of English supported the introduction of e-portfolios, most set their minds against all forms of automatic rating of content and language structure. QCA and the awarding bodies therefore began to work together to get the teaching profession on side and the provision of greater choice of assessment strategies and question types became one method by which teachers' support was gained.

## **The development of on-screen, on-demand AS/A-level examinations**

The development of e-assessments for AS/A-levels began in 2009. Given the schools' response to on-screen only syllabuses at GCSE, the unitary awarding bodies were allowed to proceed directly to on-screen, on-demand assessment of AS/A-level syllabuses. The unitary structure AS/A-level syllabuses units and the experiences of GCSE development speeded the process. While it was rare for all of an AS/A-level unit to prove assessable using a single e-question type and marking method, the grouping of e-question types into sections and the development of section banks became the norm.

One feature of QCA's requirements for AS/A-level unitisation, that initially appeared to be incompatible with the concept of on-demand e-assessment, was the requirement for the synoptic assessment. Despite pressures from schools and awarding bodies QCA remained reluctant to concede this point, partly - it was believed - because to do so would have attracted criticisms from media pundits that standards were being lowered, again! Reason prevailed and the issue was resolved by the agreement that additional synoptic assessments opportunities would be timetabled.

It is interesting to speculate what the consequences might have been had QCA decided to impose synoptic units as part of GCSE e-assessment. To have done so would have risked negating the concept of on-demand assessment and might well have deterred schools from participating in on-demand assessment of the examination.

### **The wide scale adoption of on-demand e-assessment**

By 2010 the range of e-assessed GCSE subjects provided by the unitary awarding bodies included most of the large-entry subjects and on-demand assessment had become the norm in FE and sixth form colleges. However, secondary schools remained reluctant to use on-demand assessment with more than a few subjects. The principle issues were the supply of good quality e-learning materials to support personalised study of GCSE subjects, and development of the technical, physical and human resources to support e-assessment.

### **The availability of e-learning materials**

By 2009 it was clear that the e-assessment revolution was not going to happen unless the supply of e-learning materials to support personalised learning was addressed. The approach to the development of e-learning materials adopted for the GCSE and AS/A-level examinations differed from the one used earlier for vocational qualifications. First, a wide variety of agencies were involved in the work - awarding bodies, publishers, the professional bodies for the sciences, mathematics, and geography, subject associations, groups of teachers, media organisations and even the Nuffield Foundation, all working in different combinations over two to four years. Second, all of the results from the formative and diagnostic on-screen assessments contained in the developing materials were collected and used to refine the materials. Further, studies were conducted to identify and disseminate good practice in personalised learning and on-demand e-assessment. Not only did these processes generate high quality learning materials, they also addressed many of the teachers' concerns about both developments.

By the end of 2011 schools had access to at least one e-learning scheme tailored to each of the unitary awarding bodies' on-demand GCSE subjects. All of the schemes contained:

- Within topic or module assessments for students to test their own learning;
- End of topic or module assessments for teachers to monitor student progress;
- Diagnostic end of unit assessments for teachers and students to evaluate the latter's readiness for on-demand external assessment; and by 2012,
- Remedial units that focussed on common areas of weakness.

All of these types of assessment are now supported by systems to feed the outcomes back to learners and their teachers, with personalised learning options and recommendations. The effects on teaching and school/ college organisation were profound. Instead of de-skilling teachers, as many had feared, the feedback arrangements came to focus and support their work in new ways. However, the very success of the personalised learning led to new areas of controversy.

Whether or not they were using e-learning systems, schools found that it was relatively simple to manage e-assessment for a small number of subjects. The schools were able to maintain the necessary number of workstations, they had existing ICT teaching and support staff to prepare them and the effects on the timetable were manageable. However, when e-assessment was extended across the GCSE curriculum, new arrangements had to be made. As the numbers of GCSE and A/AS-level subjects using e-assessment increased the very way that schools and colleges were organised became challenged. For example, external e-assessment became a year-round issue. A typical school or college with 250 GCSE students taking an average of 8 to 10 GCSEs, that required - say - 3 e-tested units per subject (and making an allowance of 20% for missed sessions, resits and other technical and administrative disruptions) had to manage 8000+ GCSE on-screen test presentations per year. Given that most students were likely to be ready for their assessments towards the latter part of year 11, the school might have only 100 school days on which to deliver them. This implied an average rate of 80 on-screen, on-demand sessions per day. When other qualifications such as AS/A-levels, Key Skills etc were factored in, this daily on-screen presentation effectively doubled<sup>xi</sup>.

In short, once a school tried to extend e-assessment beyond 2 or 3 subjects, muddling through was no longer an option. To do so risked reducing the resources that could be allocated to teaching, learning and the preparation of e-portfolios. Instead, schools had to provide specialist e-testing facilities, with their own workstations and staff to manage them. The personalised pattern of delivery developed with individual students attending at the e-testing suite at specified times, waiting until a workstation was free, logging-on and taking their designated test, receiving any results and/or feedback available and returning to scheduled lessons. Using a nominal school day of 6 hours and an average e-test length of 60 to 75 minutes implied that 20 to 40 students were taking e-tests at the same time. Extending e-testing beyond the typical school day and to other days in the school year became inevitable at peak times. Those students who were able to take tests earlier and later in the day were encouraged to do so and resits were often confined to evenings, Saturdays, even holidays.

E-learning schemes for the same subjects were found to make even greater demands on institutions resources. E-learning made even greater use of assessment than qualifications delivery - albeit for different purposes - and it is easy to see why internal assessment tended to remain the responsibility of subject teachers. The conclusion reached by many teachers and others was that the levels of internal and external assessment required to implement and maintain personalised learning and e-assessment were unsustainable. Many came to see the overlaps between external and internal e-assessment to be wasteful of resources. Therefore, some of the e-learning and e-assessment schemes produced for the GCSE after 2012, came to mirror the Kent Maths and GA schemes of earlier years and use common e-assessments that are being used to support students' learning and contribute to subject awards. The pattern is also developing for some internal e-assessments to be externally e-marked. Similar schemes for AS/A-level studies are under consideration at the time of writing.

Although the KS3 ICT pilots and assessments provided most schools with their first experiences of personalised assessment delivery, the many unique features that they contained meant they were limited precursors of general e-testing systems. Nevertheless, QCA was to find many other uses for the KS3 ICT net, delivery platform and functions.

## **Developing the resources for e-assessment**

The on-screen only pilots of 2006 and 2007 were predicated on the existence of school/ college ICT networks to deliver the unit assessments. Maintaining large numbers of workstations and loading all the necessary software to individual workstations was considered to be unworkable. The 2006 and 2007 pilots also showed the problems associated with institutions having to load and maintain multiple sets of software in order to deliver the assessments of different awarding bodies. In 2006 QCA tried to address the problem by specifying a common delivery system (platform) for all awarding body e-assessment systems. However, the pattern that developed after 2007 was for all the software needed for an on-screen assessment session to be downloaded to a candidate's workstation, either direct from the awarding body or the centre's hard drive, at the start of each on-screen assessment unit.

Schools also found that they had to address the resource issues associated with wide-scale e-testing and e-portfolios. Physical space proved to be the priority for the former and many institutions created dedicated e-testing facilities to deliver on-demand delivery of assessment units. The alternative to this approach was for schools to purchase time at the commercial, local college and LEA e-assessment facilities that arose in response to these needs.

Staffing the dedicated e-assessment facilities was also difficult for two reasons. Initially, there were insufficient qualified staff to manage and maintain institutions' e-delivery systems. A system of national qualifications for e-assessment delivery was accredited by QCA in 2009<sup>xii</sup> and many LEAs sponsored training courses. However, successful students tended to drift away to better paid ICT work in industry.

The main role of the ICT staff supporting e-testing became:

- Maintenance of the delivery network and workstations;
- Monitoring the download of assessment units from the awarding body to a designated workstation at the start of each student's assessment session;
- Monitoring the upload of each student's responses to the correct awarding body;
- Receipt and production of results from each assessment;
- Security, accessibility and data protection.

Examination officers and teaching staff tended to retain responsibility for:

- The overall management of e-assessment;
- Allocation of e-resources for e-testing;
- Registration of students for e-assessment;
- Programming of on-demand assessments;
- Management of invigilation;
- Issue of results and feedback to students.

## ***E-portfolios***

Although e-portfolios were widely accepted as a "good thing", their implementation on a large scale proved to be problematic. The main issues continue to be the number of available e-portfolio systems, their features and efficiency. By 2007 institutions were faced with a bewildering choice of systems devised by commercial producers, awarding bodies, some LEAs, a few school-designed systems, and even offerings by individual students. Related issues for all institutions were the comparative efficiency of different systems and the quantity of e-resources that could be made available to individual students. Increasingly, Institutions came to state in their prospectuses the number of gigabytes that they promised to allocate to each e-portfolio.

Attempts by QCA to foster (and later impose) a single system failed. The situation has improved and it is now accepted that awarding bodies can nominate "preferred" systems. However, given the numbers of unitary and vocational awarding bodies, and the differing requirements of the latter, most institutions still have to maintain several systems. Even today many staff and students have to become

familiar with more than one e-portfolio system. The features of the different portfolio systems and their efficiency remain topics for heated debate.

One area where QCA had to intervene, after some high profile failures of institutional systems and subsequent loss of students' e-portfolio work in 2007, was the imposition of standards for e-portfolio storage, retrieval and backup. This work was taken forward with the British Standards Institute and became the first of the joint standards for e-assessment systems.

In most institutions, the management of the compilation, progress monitoring, internal assessment and authentication of e-portfolios have tended to remain a faculty responsibility.

### ***E-vivas and e-presentations***

Arrangements for e-vivas and e-presentations tended to vary from school to school. The most common pattern that developed was for their management to be left to individual faculties to arrange, according to the assessment policies of the institution, especially when specialist equipment and facilities were required. However, when central facilities were to be used, such as video conferencing facilities for language orals, responsibility has remained with the examinations staff.

### ***Organisational issues***

By 2010 it had become clear that delivering the potential outcomes of personalised and e-assessment required changes to the traditional patterns of school and college organisation. The problem was not just that personalisation of learning led to students nominally working to the same timetable becoming ready for external assessment at different points in each course, but what to do with when they had completed qualifications. The solution adopted by many schools and colleges was to abandon the traditional structure of separate GCSE and AS/A-level courses and allow students to begin the latter, as they completed the former. Concurrently, new tools were needed to enable teachers to monitor students' progress, ensure that students knew when to attend for on-demand assessment and inform teachers which students to expect for e-learning sessions. Like external assessment, timetabling became a 200-day per year activity.

The prime characteristic of the successful personalised curriculum is its flexibility, not only to provide e-testing opportunities as students became ready, but also to respond to the outcomes of e-assessment and enable successful students to move their focus to other units in the same subject, to other subjects and other qualifications.

However, by 2010 it was becoming clear that some students were not profiting from the personalised curriculum and would make more measurable progress when they learnt in traditional ways. The solution to this second problem was to extend the concept of the personalised curriculum to include blends of traditional and e-learning to suit individual student's needs. Examples of how some institutions have responded to these organisational challenges are discussed in the case studies reported in Part D (below).

### ***Postscript***

Perhaps the most bizarre detail in the histories of e-assessment, e-learning and the personalised curriculum, was the ongoing hostility of one pressure group that advocated traditional teaching and examinations. From 2006 it published several papers and reports that derided all aspects of e-learning and e-assessment. The group's attention eventually focused on the use of HE units as advanced and enrichment units for students that complete A-level studies early. In 2013 the group made a complaint to the National Audit Office that the teaching of HE units by schools was illegal, citing the 1901 Cockerton Judgement<sup>xiii</sup> as a precedent. Following an investigation, the NAO ruled that, while the delivery of complete degree courses in schools would contravene the earlier judgement, the teaching of individual units did not.

## Part C: Features of the assessment system in 2014

---

The use of paper for the recording of students' responses is now confined to subjects and qualifications where manual writing skills are still deemed to be important and a small number of qualifications and levels that would be uneconomic to assess on-screen. Similar cost pressures also preclude the use of paper, section and question banks for such units even when they are assessed on-demand. At the time of writing, many hybrid assessments remain that involve combinations of on-screen or on paper delivery and response capture. Even when students still respond on-paper to external assessments, QCA expects their responses to be rendered into one of the standard electronic formats so that they can be e-marked and the codes of practice for e-moderation and feedback to students can be complied with.

Automatic rating of English remains controversial, for both written and oral responses. Despite the additional costs involved many schools (and parents) still prefer to have their students' responses for English language and literature assessed by expert markers.

Those awarding bodies that have not adopted on-demand assessment for all of their units remain under pressure from students demanding improved access to resits. QCA has resisted pressures to make early resits a requirement for all qualifications.

### **Registration for GCSE and AS/A-level qualifications**

QCA oversees a life-long record of each student's qualifications that begin when they take their first National Curriculum key stage assessment. Registration of students for external qualifications has become a two-stage process. First, institutions register their students for qualifications with the central system overseen by QCA. As each student becomes ready for e-assessment, a request for the on-demand assessment of a particular qualification unit is sent to the relevant awarding body. Delivery of the assessment unit, marking, feedback and the issue of results slips are all the responsibility of the awarding body. Attempts by the central system to manage, and later to monitor, on-demand assessments proved to be impracticable, so it is not until the institution requests certification of a student's result for a subject that the central system updates the student's records and issues the certificate.

### **Syllabus delivery media**

Schools now have some choice about how units of GCSE and AS/A-level syllabuses are structured and assessed. While all three unitary awarding bodies have on-screen, on-demand versions of all of their large entry syllabuses, on-paper versions of most GCSE units are still available. A few schools continue to select mixtures of on-paper and on-screen units, claiming that by doing so they meet personal needs. Teacher and other pressure groups continue to argue that on-paper assessment should be the preferred medium for some GCSE units.

The real choice, however, is between the on-screen syllabuses of different awarding bodies. The agreement of on-screen question types over time has meant that the mix of question types used in a particular e-assessment unit is partly a function of when it was approved by QCA. Some teachers claim that the contrast between "new" and "old" question types is impacting on teaching because different aspects of the same content are emphasised.

## Question types and associated marking methods

Essentially, six general question types are being used for on-screen qualifications delivery. However, the ways that these basic types have been implemented are legion.

1. Coded response questions (*for automatic scoring using OMR and ICR systems*);
  - a. MCQ;
  - b. MCQ equivalents;
2. Question types linked to workstation functions (*for automatic scoring*):
  - a. Highlighting;
  - b. Linking of pairs and combinations of responses;
  - c. Labelling maps and diagrams;
  - d. Selection of stated numbers of items from lists;
3. Questions requiring written (keyed) responses of different lengths (*for automatic rating of content and/ or language structure, clerical marking and expert marking*):
  - a. Short answer - words and phrases;
  - b. Short and extended writing;
  - c. Calculations;
4. Questions requiring oral responses, performance, practical work in progress (*for automatic rating of content and/ or language structure, clerical marking and expert marking*):
  - a. Oral and aural work (e-vivas);
  - b. E-presentations, via Internet and video conferencing;
5. Scenario questions that utilise a mixture of the above question types to assess a common theme or content;
6. Problem solving questions based on the use of mathematical models.

Many of the above question types were based on and looked similar to MCQ and other coded response questions. Given the dependence of all of the above question types on psychometrics to calibrate and validate the new question types, it is difficult now to understand that the use of MCQs for general qualifications was ever banned.

One form of e-testing that was believed, in 2004, to show promise was adaptive testing. Instead of expecting all students to answer the same questions, the proponents of adaptive testing claimed that tests would be more efficient if the later questions were varied in response to a student's answers to earlier ones. Adaptive testing is now understood to be most efficient when used to distinguish between students with heterogeneous levels of attainment or prior learning. Therefore, examples of the technique in external assessment are rare, but it is used widely as a diagnostic technique when students begin new e-learning modules.

## Banking of questions, section and papers

Papers for most unit assessments are now derived from question, section and paper banks. The management of these banks is outside the scope of this paper but what is of interest is:

1. The degree to which the questions, sections and papers in the banks have been psychometrically calibrated;
2. The degree of confidence that can be attached to the calibration of each on-demand paper when it is first used; and,
3. The confidence with which the results for individual units can be aggregated to make on-demand awards for complete qualifications and subjects.

While on-demand assessment was largely limited to vocational and Basic/ Key Skills qualifications, little concern was expressed about the calibration, confidence and aggregation issues. Passes in the individual units and in the overall qualification tended to be all that was required. However, when on-demand GCSE and AS/A-level grades were introduced, there was immediate controversy.

The teaching profession and educational media were prepared to accept that grades derived from unit awards, that were in turn derived from fully calibrated banks, would be secure. The first issue was that the only question types that could be "fully pre-calibrated" at the time were multiple-choice questions, which were banned anyway until 2008. As the questions increased in length and complexity, they tended not to be less pre-tested for reasons of cost, risk of disclosure and practical issues to do with the compilation of unit assessments. Therefore, the expected systems of pre-hoc calibration were not realised. Instead, a new form of post-hoc calibration developed under which a unit is considered to be "partly calibrated" until its calibration could be confirmed using adequate numbers of students' results. Even then, QCA requires that the calibration of each banked question, section or paper remains under review and earlier results are re-evaluated when recalibration proves to be necessary.

The second issue was how, using partly calibrated unit assessments, adequate information could be fed back to students quickly? Listing, with reasons, what a student got wrong and providing relevant advice was the easy part. The difficulty was how to report a level of performance in terms of putative GCSE and AS/A-level grades? The very real concern was that a partly calibrated unit assessment might prove to be more difficult (or easier) than expected, after a student had taken it. Such variations in standard have been a feature of all live examinations since the first university matriculations of the early 1830s, however thorough the pre-testing and paper moderation procedures became. Up to 1960 marks were scaled so that comparisons could be made across years; from 1960 awarding bodies used grades and simply changed the mark boundaries for each grade at the end of each examination cycle.

The particular concern of students, teachers and parents (and therefore QCA) has been that a student, who was awarded passes for each on-demand unit might not pass overall because one or more unit has been re-calibrated. In practice, such fears have proved to be largely groundless. The awarding bodies are conservative with their initial calibrations and provide overlapping mark ranges for each grade. The overlaps in the mark ranges are sufficiently wide to remain valid even if a unit is re-calibrated. They are a guide to the final level but little else. As a consequence of these strategies students are more likely to be surprised by the grades they received than to be disappointed. (See also the section below on on-demand awards.)

### **The hierarchy of assessments used in e-learning materials and external assessments**

Given the ease with which assessments, particularly MCQs tests, can be generated from banks, many e-learning courses contain assessments that contribute directly to the award of qualifications. Many awarding bodies have reasoned that formal examination conditions are unnecessary for many uses of assessment. Instead, the degree of rigour associated with an assessment should reflect its purpose and level. As a consequence, QCA now recognises that a blend of e-learning and external assessments is appropriate for many qualifications. Examples include:

1. Self-assessments taken as part of student's personalised learning - anywhere and at any time;
2. Assessments taken in trust conditions - at home, in libraries etc;
3. Assessments taken during class time;
4. Assessments taken under external examination conditions.

### **Supporting and monitoring the delivery of e-assessments**

Given the number of early e-assessment sessions that were disrupted by technical and managerial failures in institutions QCA has cooperated with the awarding bodies and the British Standards Institute to regulate e-assessment delivery at the institutional level. Series of joint standards for all aspects of e-assessment have been produced and are revised regularly as practical case law accumulates.

The awarding bodies provide training in their systems and 24/7 help lines for teachers and technicians. The awarding bodies, QCA, LEAs and OFSTED also conduct technical inspections and audits of institutions' systems, before, during and after e-assessment sessions.

Media opinion continues to demand even tougher approaches. The option of an awarding body or QCA establishing a link into an institutions' ICT network, when it appears to be incapable of managing e-assessment directly, has been discussed but has always been rejected as unacceptable. Instead, when institutions are unable to provide e-assessment services to their students on a temporary basis, due to a short-term deficiency of technical or human resources, they are referred to a local commercial test centre. Were a longer term inability of an institution to provide the necessary e-assessment resources ever to be identified it would lead, inevitably, to de-accreditation by the relevant awarding bodies.

The debate, about whether repeated failure of e-assessment sessions and loss of e-portfolio materials should be considered as malpractice, continues. However, QCA codes do require the maintenance of an e-assessment log and the reporting of all interruptions to external e-assessment sessions, of data loss and data corruption (to the awarding body in the first instance).

### **Supply of e-markers**

Most responses to KS2/3, GCSE, and AS/A-level questions are now marked automatically using OMR/ICR scoring or the automatic rating of content and language structure. Clerical and expert marking is used with large entry qualifications to develop and audit the automatic marking systems and for selected purposes such as some marking of calculations, extended writing, practical work in progress, performing arts presentations, oral and aural work.

Teachers' interaction with the external assessment process has changed in two ways. First, visits to institutions for examination, moderation, and verification processes have largely given way direct links between the assessor and student via the Internet and video conferencing. Current QCA regulations required that, however an assessment unit is marked, the students' responses should be recorded in an approved electronic format, with full audit trails of the marking and moderation. Second, more teachers now contribute to the creation and review of on-demand questions and e-learning units than work as expert markers.

The awarding bodies continue to employ teams of markers but on permanent (200 days per year) rather than seasonal, bases. Clerical markers now work almost completely in awarding body centres. For convenience, the expert markers employed to maintain and audit the marking carried out by automatic system and clerical teams also tend to work in awarding body premises but most work via the QCA and awarding body Intranets. Individuals based in the UK and overseas continue to be employed as expert markers and some schools have contracted to provide year-round marking services.

The stimuli for most of the new marking arrangements were, of course, cost and speed of reporting. Major factors in their design these systems have also been the control of the standards and timeliness of the marking processes.

### **Awarding of on-demand qualifications**

The traditional model of post-assessment awards has not been completely replaced by the partial calibration of questions, sections and papers, followed by confirmation of the calibration using large samples of students and any necessary recalibration. The main issue is the aggregation of unit results into overall awards. Awarding bodies are expected to keep the calibration of all banked questions, sections and papers under review. When evidence is found of change in the standard of an assessment unit, recalibration is undertaken and all incomplete sets of students' results that might be affected are checked and, if necessary, updated.

As early as 2006, it was recognised by QCA and the awarding bodies that identical sets of raw marks for different units would not necessarily lead to the same awards. Further, marks awarded for units might even change and there were fears about how students might react if one or more of their unit assessments were to be re-calibrated. Fortunately, most teachers and even many parents have come to understand that re-calibration of some units may be required to achieve fair results for all. Nevertheless, it is QCA's policy that initial calibration of each e-assessment unit should be sufficient for re-calibration to be the exception and not the norm.

The solution adopted by most awarding bodies - vocational and unitary - is to provide a web page for each student and his/her teacher so that they can monitor the accumulating results. All resit results are recorded and the best result for each unit is used for aggregation purposes. QCA considers that the process of recalibration should be transparent and that it is good practice (but not mandatory) for any recalibration of units to be indicated on each student's web page.

As the students complete all of their unit assessments, they can visit their web page and use a routine to indicate the grade that they will receive, given their current results. This projected grade is then discussed with parents and teachers for the latter to action the certification process. (Students are not able to do this.) Some versions of the routine also indicate any particularly low unit results and suggest how they should be improved for a desired overall result. Similarly, when students have completed all but one of their unit assessments, they can use another routine to indicate the level of performance required for a particular grade outcome - within of course a range of possible grades.

Part of the feedback to students following each external unit assessment is a notification of results slip, provided by the awarding body. Other than the mark obtained, these slips contain little information of real value. Nevertheless, they continue to be used at job and college interviews. Schools, students and parents all react vigorously whenever it is suggested that the slips should be discontinued.

QCA and the awarding bodies work together to maintain the comparability of individual unit assessments, and complete qualifications at their 2004 levels, across delivery media, across awarding bodies and across subjects. Evaluations of all forms of comparability are dependent upon the quality of the data derived from e-assessments. However, as time has passed and e-assessment has developed, maintenance of the first form of comparability has become increasingly subjective due to the limited quantities of relevant results data available from 2004.

## Part D: The case studies

---

To evaluate the impact of e-learning and on-screen, on-demand assessment of the GCSE and AS/A-levels, a series of questionnaire studies and visits were conducted. From the responses received up to September 2013, four institutions were selected for in-depth study:

1. A large 11 to 19 city comprehensive school;
2. A further education college in a new town;
3. An 11 to 19 urban grammar school;
4. An 11 to 16 comprehensive school in a rural area.

The first two institutions have adopted e-assessment and e-learning fully, and are considered to have developed good practice in the delivery of personalised curricula to their students. The second two have adopted e-assessment only partially.

### Case study No 1: The large 11 to 19 city comprehensive high school.

This school embraced all aspects of e-assessment from the start, but critically. It gained some early experience of e-assessment from on-screen, on-demand delivery of the Key Skills tests and was a pilot school for the KS3 ICT project from 2004. The school was also involved in some of the first pilots of on-demand only GCSE examinations in 2006, which it delivered on an individual basis. While difficulties were encountered with the GCSE pilots, the school learnt from them and:

- Designated an e-assessment room, later extended to a suite of rooms, with workstations that were not used for other purposes;
- To assist e-resource management, separate networks were created for e-assessment, e-learning, e-portfolio compilation, teaching support, and school administration.

The school realised from the first that the effects of e-assessment and e-learning on school organisation would be profound. The development of the current personalised learning and assessment curriculum was achieved in several steps. A common feature of the changes has been the use of the feedback from external (e-assessment) and internal (e-learning) assessments to inform the ways that teaching, school organisation and personalised courses have developed.

Initially, the school saw e-assessment as offering opportunities for students to reduce the pressures of GCSE assessment by spreading their unit assessments over a longer time period. Staff reasoned that taking some subjects early would allow students to devote more time to the subjects that they found difficult or begin one or more AS courses early. The staff also believed that less able students would be motivated by early success in some GCSE units. Rather than selecting a core of subjects for on-demand assessment, the school decided that each student should be allowed to select one or more subjects to take early. An open policy on resit assessments was adopted, but parents were required to pay for the second (and any subsequent) resits and had to agree to all resits being taken out of normal schools hours.

After 2008 sufficient quantities of GCSE e-learning materials were becoming available for the school to change the organisation of years 10 and 11 to enable a significant proportion of students to complete most or all of their GCSE examinations early. The teachers remained concerned that the use of e-learning and on-demand e-assessment required students to be mature enough to take major responsibilities for their own learning. Therefore, the school created two streams, the first based on personalised learning and assessment, and the second using more traditional teaching methods but with most students taking some subjects early to reduce their overall assessment load. In effect the school created a "school-within-a-school" à la Borland (1989)<sup>xiv</sup>

During 2009 and 2010, students in the personalised learning stream were able to complete most or all of their GCSE examinations early and:

- Devote the time to the study of other subjects;
- Study related subjects such as statistics following passes in mathematics, and the separate sciences following single or double awards in (integrated) science;
- Study enrichment and advanced units that extended the study of subjects that they planned to study at AS/A-level. (The school is one of the institutions that have lobbied QCA to allow successful study of enrichment and advanced units to enhance GCSE grades);
- Study for other qualifications;
- Start the study of advanced courses (AS/A-levels) early.

In 2011 students were able to take the first on-demand A-level and many completed subjects early. Unfortunately, most institutions of higher education were very unwilling to take students that were not adults. Therefore, the options available to these students were limited to:

- Taking further AS and A-levels or vocational qualifications to broaden and advance their sixth form studies. The school advocated several cross-curricular options, such as languages for science and technology students, and mathematics or sciences for arts and humanities students. Despite some very encouraging results, these courses have not become popular;
- Applying to one of the universities that would take students early;
- Leaving school early to take their gap year;
- Leaving school to enter employment (not encouraged);
- Sampling the study of their preferred subjects at higher education level by studying Open University units and courses provided by local universities. (The school organised a campaign to fund this initiative.)

From 2012 the school adopted its current pattern of organisation for years 10 to 13. Experience had shown that dividing students into independent and traditional learning streams was not meeting all students' needs and the school's examination results were reflecting this. Concerns were expressed that students who were mature enough to profit from the personalised learning for most of their subjects might have improved their grades by studying some subjects in the traditional stream. Similarly, many students in the traditional stream showed that they would have been able to profit from personalised learning for some of their subjects as their GCSE courses progressed and they matured.

The schools' solution is to offer all students access to the personalised learning and assessment opportunities at GCSE and A/AS-levels in the "school-within-a-school" for as many of their subjects as they are able to manage. The typical pattern is for students to study some subjects in each stream, with GCSE and AS/A-level studies overlapping in years 11 and 12. Experiments also began with "blended courses" that used combinations of e-learning and more traditional methods to deliver the personalised programmes.

As the full implications and opportunities of the personalised curriculum became clear, the roles of the teachers in the school gradually changed from deliverers of the curriculum to managers of personalised learning in their subject areas. Even within the personalised curriculum some whole class activities are retained but their nature varies with the subject. Those subjects with a practical, oral, e-portfolio and performance components tend to designate separate sessions for these forms of learning. Team teaching has developed whenever parallel groups for the same subjects can be timetabled.

Teachers' class-based work now includes occasional introductions to new topics and modules, work with individuals and groups that have difficulties with aspects of their current e-learning unit, managing e-learning and assessments, monitoring the progress of individual students, and using feedback from internal and external assessments to counsel students. The blended courses are popular with the teachers but their effectiveness has not yet been evaluated.

Evidence of the positive effects of the personalised learning curriculum has come from several sources. First, it was well received by the parents of students that took assessments early (and passed). Many parents and students have realised that obtaining the highest possible GCSE grades matters less for university entrance when there are early AS and A-level unit results to report. A few students, however, had to be dissuaded from undertaking repeated resits in order to improve already adequate GCSE grades. Certainly, the schools' overall academic results have improved steadily. Most importantly, positive reports about the students' maturity, motivation and independence have been received from universities and employers.

The latest initiative of the school has been to form a link with its local university to provide access for students in years 12 and 13 to selected university lectures and faculty events.

### **Case study No 2: The college of further education**

With three important exceptions, the experiences of the college parallel those of the first case study school. Given the earlier availability of e-learning materials from some vocational awarding bodies, the college was able to introduce an independent learning and assessment regime for some qualifications as early as 2005, including Basic Skills. The regime was extended to other qualifications as materials became available and to its general (academic) courses after 2009.

Also from 2005 the college instituted a diagnostic programme for all its September 2005 entrants. Key skills tests drawn from QCA's website were used to evaluate each entrant's abilities in language, numeracy and ICT. Those that passed at Level 1 in a subject were offered the Level 2 assessment the next day and so on. By the end of a student's first week, the college was able to timetable a general course that addressed the student's strengths and weaknesses. Similar diagnostic programmes running concurrently were used to determine the starting points for each qualification studied.

The most innovative element of the programme came in later in the 2005-2006 academic year when the college instituted an enrolment day every Monday. New students spend their first three to five days in the enrolment unit working through the relevant general and vocational diagnostic tests before the creation of their personal e-learning timetable, with target dates for their initial e-assessments. Students interviewed said that they welcome the positive aspects of the programme and its response to their personal needs.

### **Case study No 3: The 11 to 19, urban grammar school**

The school is noted for its traditional outlook and its examination results at GCSE and AS/A-level. The headteacher and staff were not committed to the concept of on-screen, on-demand assessment of the GCSE but felt that they had to be involved for fear that paper-based GCSE examinations might be discontinued for some subjects after 2009. The school had been invited to pilot the first KS3 ICT assessments in 2004 but had declined. In 2005 the school had participated but had found it difficult to arrange all of the assessments during designated ICT periods.

The school had been one of the few to pilot the on-screen only GCSE examinations in 2006 and 2007 but, to keep timetable disruption to the absolute minimum, had not used the individual delivery model of delivery, despite the recommendations of its awarding body, because it was felt that the importance of the GCSE examination would be diminished. The Deputy Head (Academic Studies), who was also the examinations officer, stressed the importance of students starting and finishing GCSE examinations together. Personalised delivery of on-screen tests, he suggested, would lead to students failing to

check their answers thoroughly and some might feel that they were under pressure to complete their unit quickly. Instead, the school had tried to replicate its normal GCSE examination procedures and, as an alternative to the sports hall, had attempted the simultaneous delivery of on-screen units to students working in classrooms and laboratories that were connected to the school's ICT network.

The school adopted a different policy to the delivery of the Key Skills qualifications, which it saw as alternatives to the GCSE for students with poor grades or as broadening qualifications for mainstream AS/A-level students. As a consequence, the school decided to offer Key Skills qualifications on-demand, using individual delivery, when Level 3 tests became available on-screen in 2007.

Many of the early on-screen GCSE assessment sessions failed in part or completely. In the subsequent discussions with the awarding body it was found that too many of the workstations and laptops used had insufficient free hard disk space, several were below the awarding body's recommended specification, the network was being operated at near capacity and a software update that provided special language fonts had not been loaded to all relevant workstations.

Given these problems the school decided to discontinue on-screen assessment for 2008 and 2009 but made renewed attempts in 2010, using full on-screen, on-demand examinations in three GCSE subjects. While these attempts were successful, the school decided not to increase the number of subjects for a range of reasons, including:

- Concerns about what the school was expected to do with students who completed a majority of their GCSE subjects early;
- Belief that on-demand assessment would disrupt classroom teaching and school organisation;
- Belief that the availability of resits would diminish the status of the GCSE.

In short, the school was concerned that the introduction of on-demand GCSE examinations challenged the very factors that led to its previous record of success. Therefore, the on-screen, on-demand GCSE assessments that were taken early were used to reinforce the school's academic policy. Examples of GCSE subjects taken early included:

- Integrated science at the end of year 10. Students with adequate grades are then able to study one or more of the separate sciences;
- Mathematics. An accelerated group now begins AS level in year 11;
- French. So that specialist linguists can devote extra time to their second (and any third) modern foreign language.

One aspect of e-assessment that the school has embraced fully is the use of e-portfolios. From 2008 all GCSE and A/AS-level course work and Key Skills portfolios have been compiled and submitted using e-portfolios. The school is particularly pleased that original art and DT work can also be photographed for external assessment and submitted as e-portfolios as the originals remain in school for open evenings. The school does not use the e-viva and e-presentation systems and prefers to have visiting examiners, even at the extra cost.

Results for this school have not risen as steadily as those of others. The school attributes this in part to over-use of resit opportunities by other institutions.

#### **Case study No 4: The rural 11 to 16 comprehensive school**

This school failed to use either the early on-screen, on-demand Key Skills tests or the pilots of the KS3 ICT assessments until both became mandatory. When it did use the latter the experience was negative, so the school initially saw the claimed advantages of on-screen, on-demand assessment of the GCSE as applying to other institutions. For example, it saw no value in having students take GCSE unit assessments early in their GCSE courses, unless they could complete sufficient subjects to be able to transfer to the local sixth form college in a nearby town or the college of further education, about the same distance in the opposite direction. The sharing of advanced and enrichment courses with other schools was precluded by the times that students would spend travelling.

After some early failures with live on-screen only units in 2007, from 2008 the school focused on enabling its students to take their GCSE units as far as possible on-demand in the second half of year 11. This, the school reasoned, reduced the overall burden of external assessment while allowing adequate time for essential resits. Exceptional students were able to take some units on-demand during the second half of year 10 and the first half of year 11. No enrichment or advanced units were provided to successful students. They were expected to join other classes so that they could devote extra time to the study of outstanding subjects or bring other GCSE work to the subjects that they had passed and work silently at the back.

## **Part E: Outcomes, implications and projections**

---

### **Changes to the nature of the GCSE and AS/A-level examinations**

The introduction of on-screen and on-demand GCSE and AS/A-level examinations, particularly the latter, changed the nature of both sets of qualifications and blurred the distinction between them. Students are able to study GCSE and AS/A-level courses at their own rates. Using the enrichment and advanced units they are able to study both to greater depth and breadth. Had systems for rewarding students that study these extra units with GCSE and AS/A-level credits been agreed, it is likely that the take up would have been larger.

The effects of personalised assessment and learning have been even more profound. The consequences of students being able to study at their own rates, take GCSE and AS/A-level units early, receive detailed feedback on their performances and take resits on-demand have all meant that they are achieving higher grades in greater numbers of subjects. Many students are now starting AS/A-level courses early. Attempts by traditionalists to suppress the study of HE units in schools and colleges were unsuccessful so the use of Open University and other higher education units is now common.

E-assessment and e-learning have proved to be popular with most stakeholder groups. Surveys of students in schools that have adopted the personalised learning and assessments models, their parents and teachers, have all expressed very positive reactions to the changes. In contrast, the responses of students in schools that have not adopted, or have partially adopted, e-assessment and e-learning systems indicate that they feel left out of an important educational development.

Despite every effort by QCA and the unitary awarding bodies to maintain the standard of individual GCSE and AS/A-level units, there are still those that equate the higher grades obtained with lowered standards. Perceptions that the selective function of the examinations has been diminished are common but views on whether this is a "good" or "bad" thing vary.

### **QCA and the regulation of e-assessment**

E-assessment has changed both the focus for the regulation of qualifications in the UK and the relationships between QCA, the awarding bodies and institutions. With respect to the first, QCA can be said to have both lost and gained influence. While its previous domination of the detail of GCSE and AS/A-level syllabuses and assessment structures is reduced, it directly manages the student registration and record systems. While the attempt to create a priori principles and codes of practice for e-assessment proved to be abortive as e-assessment case law developed, QCA's development (with BSI) and enforcement of technical codes for e-assessment is widely acknowledged. Many of the tensions between QCA and awarding bodies are history. Instead, the regulators and the regulated now work as partners in a common endeavour.

There are still issues to resolve. Discussions about the numbers of resits that students should be allowed to take, credits for enrichment and advanced GCSE and AS/A-level units and when students should "cash-in" their unit results for certification all continue. A decision is outstanding on whether repeated technical failure in a school's delivery of e-assessments - due to failure to maintain the network and/or workstations, apply pre-assessment checks, or load necessary software - constitutes malpractice. Similarly, awarding body responsibilities when new e-assessment software fails, continue to be dealt with on an individual basis.

QCA has not only encouraged the awarding bodies to innovate, it has encouraged schools to experiment with new ways of organising the delivery of the GCSE and A/AS-level curricula. QCA, OFSTED and LEA personnel have cooperated to identify and disseminate good practice in e-assessment and e-learning using the QCA Innovation leaflets.

## **Implications for stakeholders**

### ***Awarding bodies***

The awarding bodies of 2014 are barely recognisable as the same organisations from 2004. Instead of being a seasonal activity, qualifications delivery takes place 200 to 300 days per year, and potentially 24 hours per day. To support this level of assessment activity, questions are being drafted, moderated, calibrated, banked, delivered, recalibrated and eventually refreshed or discarded on a continuous basis.

The marking of assessment units has also changed beyond recognition. Instead of the employment of large numbers of markers for 40 to 100 hours, spread over a few weeks per year, student's responses to most questions are marked automatically using updated versions of OMR and ICR systems, plus automatic rating of content or language structure. Many of the responses to middle and higher level questions, that cannot be marked automatically, are marked clerically. Expert markers are still employed but in small numbers and on a continuous basis. Above all else, paper based assessments are restricted to a small number of low volume qualifications.

Marking, like examining, has become internationalised. Once an image of a candidate's response has been captured, it does not matter which country it is marked in, just that it is marked reliably. By 2005, issues of marker supply and cost meant that overseas markers were already being employed. The awarding bodies also found that to employ some markers from other countries, especially if they might mark UK students' responses, encouraged entries from those countries.

For many reasons, examiners' visits to school and colleges to conduct language orals and assess performance work in the performing arts are greatly reduced. Instead, most direct contacts between students and examiners are managed by the awarding bodies, with the student/ examiner interface achieved using standard communication media. Although it contradicts the spirit of test-when-ready assessment, some schools still prefer to use, and pay the additional costs associated with, travelling oral examiners. Even then QCA requires that the sessions are recorded. In both schools and colleges, the processes and products of practical work are also examined remotely, using video cameras in the laboratory or workshop that are controlled by the examiner. All such links between examiners and students are recorded for moderation, compliance and legal reasons.

As in 2004, the unitary and other large awarding bodies operate at the frontier of ICT. One big change over the last ten years has been the extent to which the awarding bodies now communicate directly with students. All awarding bodies provide students with large quantities on feedback on their performance and some enable the students to select the form in which they wish to receive it. Many awarding bodies now provide registered students with personal web pages for them to see how their results are accumulating and the probable grades that different levels of performance will produce. Some awarding bodies provide e-learning and other support materials.

The biggest problem for the awarding bodies over the last five years has been the development of software and training to assist schools and colleges with the management of their e-assessment systems. This work continues.

### ***Schools and colleges***

The most profound changes that e-assessment has brought are to schools and colleges. The successful implementation of personalisation of learning and assessment has blurred the distinctions between KS3, GCSE, AS/A-level and HE courses. Most schools and colleges have adopted continuous learning schemes for years 11 to 13, or even 10 to 13. External assessment is now a daily responsibility.

In retrospect, it has been the schools and colleges that have dictated the pace of e-assessment developments. Their unwillingness (whether on the grounds of inadequate resources or academic choice) to participate in the non-live pilots of on-screen only GCSE units, slowed initial progress, but speeded the change to full on-screen, on-demand assessment. Full implementation of personalised assessment had to await the provision of e-learning resources and the development of new ways of managing course delivery.

Schools and colleges have profited from these changes by the introduction of e-assessment in four main ways:

1. Students now take some or all of their GCSE and AS/A-level subjects before the final term of Year 11. This means that they are able to spread their assessment load over a longer period of time. It also means that they can devote the time saved to the study of other subjects, study enrichment and advanced units for the same subject and level, or begin advanced courses early;
2. The curriculum of most schools and colleges has also broadened and deepened to include new qualifications. Some students have also been able to sample the study of their preferred subjects at the higher education level via OU units and links with local universities;
3. The quality of the feedback provided following e-assessment - both internal and external - has informed the ways that teaching structures, school organisation and personalised courses have developed. Schools and colleges that have succeeded in managing this feedback have seen their examination results rise steadily;
4. For the first time for several decades, schools and colleges have real choices about how units of GCSE and AS/A-level course are to be assessed. Not only do they have the option of on-screen and on-paper delivery of unit assessments for most subjects; the different timetables to which the unitary awarding bodies have worked have also produced variety in the question types used to assess the same subject content.

The personalisation of learning and assessment has also brought new responsibilities and issues. Formal teaching now tends to be restricted to the introduction of new topics and to subject groups that need extra support to be able to profit from e-learning systems. Instead, most teaching takes place on a personal or small group basis as issues arise from the e-learning programmes.

New systems have been developed for teachers to monitor their students' progress through their e-learning programmes and the compilation of e-portfolios. Similarly, the production of student timetables, once an annual activity, has become a continuous one as students complete units and courses, or require remedial support to prepare for resits.

One issue with the personalisation of learning has been the integration of practical work in the sciences and technology with the e-learning programmes. This has not been a problem for schools and colleges that were used to managing practical work on circus or team teaching bases; practical activities are integrated at appropriate points. However, concerns have been expressed by universities and employers that the simulations and models, that drive many e-learning activities, are being used as alternatives to traditional, hands-on, practical work.

Almost all schools - certainly all colleges - have created e-assessment suites that function during most days that the institution is open, and out of hours at peak assessment times. Managing e-testing suites has been a problem for many schools and the supply of qualified ICT technicians remains an issue. Colleges have not tended to have the same problems. The expectation that all institutions are answerable to the awarding bodies and QCA for technical mismanagement of external assessment sessions, and lost e-portfolio work, remains very unpopular.

While many teachers have complained about the changes that e-assessment and e-learning have made to their work, most speak favourably about the resultant broadening of the curriculum and the blurring of the boundaries between GCSE and AS/A-levels and HE. The profession universally acclaimed the rejection in the High Courts of the demand, by one prominent pressure group, for the reapplication of the Cockerton Judgement.

### ***Higher education***

Institutions of higher education have also gained from the use of e-assessment and e-learning. Students that have progressed through a personalised curriculum are characterised by the depth and breadth of their learning within individual subjects and across their whole curriculum, and by their abilities to work independently. Admission tutors no longer have to rely on GCSE results and estimates of AS/A-level performance. Instead, they are able to base their decisions on AS/A-level units already achieved. In many cases students seeking admission have completed their core 18+ studies, a significant proportion are studying further AS/A-level and vocational qualifications to support their proposed HE studies, and some are even studying units of HE courses provided by the Open University and others. Increasingly, applicants will have been involved in school/ university partnerships that will have led to them attending lectures and other activities.

The universities have also gained from the e-learning and e-assessment initiatives in other ways. The opportunities for income generation by assisting QCA and the awarding bodies with the development, calibration, evaluation and general research aside, the transition between school (or college) and university has become blurred. The various school-university partnership schemes, the provision of HE enrichment units, and less prescriptive regulation, have served to reawaken universities' interest in AS/A-levels and their links to HE courses.

Finally, the experimental schemes being developed by some universities and/ or employers to enable students to carry forward HE credits gained at school or college into their university careers are to be encouraged.

### ***Implication for students***

All students have gained something from the introduction of personalised learning and assessment. Students' ability to learn at their own pace and take qualification units when they are ready, the quality of the feedback provided by awarding bodies together with the security of early results, and the confidence that comes from the accumulation of unit passes are all obvious examples that have been universally welcomed by students. The overall amount of gain from personalised learning and assessment, however, has depended upon the extent to which students' schools or colleges have adopted e-assessment and e-learning, and the maturity with which students can use the opportunities that these techniques provide.

Students have also welcomed e-portfolios, but awarding body restrictions on the size of the files submitted and the range of media that can be used continue to annoy many. E-vivas and e-presentations have caused tensions when students have not been able to become familiar with the delivery systems in advance of their first external assessment.

The broadening of the curriculum and the opportunities for study outside the confines of GCSE and AS/A-level syllabuses have not always been understood by students and their parents. If offered the choice, both have tended to prefer advanced over broadening units.

### ***Parents***

In general, parents have supported the concepts of learning at the student's pace and "test when ready", particularly when the latter means that their sons and daughters will be taking some GCSE and AS/A-level units early. The possibility that the most able and hard working students can start HE units while still at school or college has proved to be very attractive.

On the downside, many examples have appeared of parents who have tried to over-pressure their sons and daughters to take some or all their GCSE and A/AS-level units too early. Further, some parental reviews of the feedback provided by the awarding bodies have been shown to lower rather than raise motivation. Instances of very successful students being over criticised for the errors that they did make in assessed units have been too common. Many school are trying to address these problems by guiding parents on how to use the feedback information.

### Exam on Demand Assessment Advisory Group

September 2005

### **Acknowledgements**

---

The contributions to this paper made by members of the ExoD Assessment Advisory Group and others are acknowledged. Particular thanks go to: Dr David Shipstone (formerly of Nottingham University, Department of Education), Jeff Ross (Broxbourne Solutions) and Bridget Kingdon (Henlow Middle School).

## Endnotes

- i QCA (2004), A proposed blueprint for delivering e-assessment, see <http://www.qca.org.uk/7192>.
- ii The Advanced and Advanced Supplementary Levels of the General Certificate of Education underwent several revisions during the period considered in their paper. For convenience the examination will be referred to as "A-level" up to 1988 and "AS/A-levels" thereafter.
- iii CSE = Certificate of Secondary Education examination which was targeted at about half of the students that were not expected to achieve pass grades in the Ordinary Level of the General Certificate of Education. The CSE examinations were managed by regional boards, based on LEA areas.
- iv For brevity, the name QCA is used for the current organisation, its predecessors and any successors.
- v Exam on Demand.
- vi Smith A (Chair) (2004), Making Mathematics Count, <http://www.mathsinquiry.org.uk/>
- vii Tomlinson M (Chair) (2004), 14-19 Curriculum and qualifications reform: final report of the Working Group on 14-19 Reform, DfES Publications (ref DfE-0976-2004), Nottingham and [www.14-19reform.gov.uk](http://www.14-19reform.gov.uk).
- viii See for example, ExoD (2005A), A response to the Tomlinson Report by The Exam on Demand Assessment Advisory Group (Occasional Paper No. 1), Exam on Demand Ltd, Hellingly.
- ix See ExoD (2005B), A supplement to ExoD Occasional Paper No. 1, the response to the Tomlinson Report by The Exam on Demand Assessment Advisory Group, Exam on Demand Ltd, Hellingly.
- x QCA (2004), On-screen delivery of qualifications, the Basic and Key Skills experience, <http://www.qca.org.uk/6993>.
- xi The calculation is based on:
  - An AS/A cohort of 150 students each taking the equivalent of 3 A-level subjects, each of which consisted of 6 units and up to - say - 8 individual tests. With an allowance for resits etc this implies over 4300 on-screen test presentations;
  - A total of 250 x 9 (3 subjects x 3 levels) Key Skills e-test presentations per year spread across years, 9 to 13;
  - An estimated 2000 e-test presentation for other qualifications.
- xii SQA - the Scottish Qualifications Authority - established a system of ICT qualifications for school staff prior to 2004.
- xiii By the end of the nineteenth century some elementary schools in major cities were teaching University of London degree courses. In 1899 the district auditor for Birmingham ruled that the practice was a misuse of public funding and should cease. His decision was upheld after appeals by the local school board in 1901 and 1902. The ruling is recorded in the history of education as the Cockerton Judgement.
- xiv Borland J H (1989), Planning and implementing programs for the gifted, New York: Teachers College Press.