



Qualifications and
Curriculum Authority

The basic and key skills (BKS) e-assessment experience report

March 2004

Contents

Executive summary	3
Introduction	4
Background.....	4
Methodology	6
Use of pilot studies	6
Feasibility studies of on-screen testing – January 2002	7
Pilot studies	8
Non-live and live pilots – May 2002	8
Non-live, live and extended live pilots – November 2002 onwards.....	8
Extended live pilots – May 2003	9
Issues relating to the participating awarding bodies' IT-based delivery systems	11
Summary of the pilots – January 2002 to date	16
Additional findings	16
An evaluation of the key features of IT-based delivery systems used in the BKS pilots.....	17
Table 1: Key features of the on-screen IT-based delivery models used in the BKS tests....	19
Table 2: A comparison and evaluation of the main on-screen delivery models.....	20
Evaluation of the project methodology	21
Conclusions and recommendations	22
Awarding body issues.....	22
Centre-based issues.....	23
Issues concerning the supply of BKS test material	23
Continuing pilot status	24
Appendix 1: The on-demand criteria	25
External assessment of level 1 and 2 key skills, adult literacy and numeracy: advisory criteria for use of the test bank.....	25
Awarding bodies' responsibilities prior to gaining access to the test bank.....	25
On-screen delivery	26

Executive summary

The aims of the project were to:

- summarise developments in the on-screen delivery of the basic and key skills (BKS) qualifications
- highlight issues with the participating awarding body's delivery systems
- identify good practice in on-screen and on-demand delivery of BKS tests.

The methodology used for the project was ongoing discussions with awarding bodies and undertaking evaluation visits.

The development work with each awarding body was structured through series of non-live, live and extended live pilots.

Five awarding bodies have developed robust delivery systems and have completed series of extended live pilots. Six further awarding bodies are developing systems. Their degree of completeness varies from the discussion and/or planning stage to pilot delivery of BKS and other qualification units (exam-on-demand), with others at intermediate stages.

A core of good practice in on-screen and on-demand assessment delivery has been identified and is being used to advise awarding bodies and service providers developing on-screen delivery systems for units of other qualifications.

An advice booklet to accompany the release of the new statutory guidelines for qualifications has been developed.

The project has been criticised by a minority of stakeholders and awarding bodies for being too prescriptive in its advice and thereby potentially creating risks of inhibiting creativity in on-screen delivery of qualification units. The project team and the BKS Project Board have rejected these criticisms. Their reasons include a comparison between the good practice identified and BSI standard 7988, and the technical diversity of the developing on-screen delivery models produced by different awarding bodies.

Conclusions that have been drawn include:

- the use of common BKS tests at levels 1 and 2 on screen requires a measure of uniformity in the on-screen delivery systems used by different awarding bodies
- good practice in on-screen assessment delivery and compliance with the BSI standard may require the awarding bodies to ensure some uniformity in the delivery systems for other qualification units.

Recommendations have been made concerning:

- awarding body issues
- centres' management of on-screen assessment
- the supply of BKS test materials
- the continuing pilot status of BKS on-screen tests delivery.

Introduction

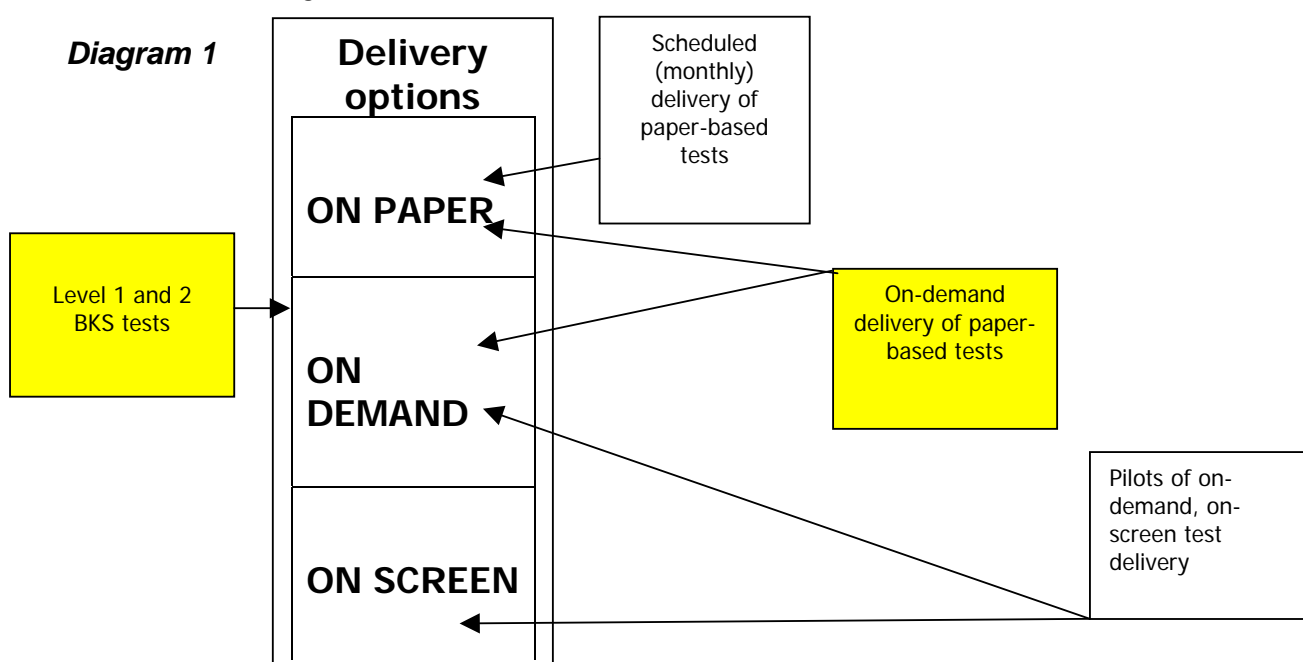
The purposes of this report are to:

- provide a summary of the current developments in the on-screen and on-demand delivery of basic and key skills (BKS) qualifications
- highlight issues relating to the participating awarding bodies' on-screen delivery systems for consideration by the regulators
- identify a core of developing good practice in the on-screen delivery of BKS qualifications to be shared with awarding bodies that are beginning the development of on-screen delivery systems for BKS and other qualifications.

Background

The BKS research studies project arose following the unsuccessful attempt, by the regulatory authorities, to appoint an agency to develop an item/test bank to deliver the BKS tests to awarding bodies. Over the past two years, staff from the e-assessment team in QCA's assessment policy and development programme (APDP) team (previously the new projects team) have been facilitating the work of BKS accredited awarding bodies to develop IT-based delivery systems to deliver the level 1 and level 2 BKS qualifications on screen.

The work is part of a wider system for making the BKS tests available on demand, whether on paper or on screen. All awarding bodies accredited to offer the BKS qualifications may deliver the external assessment component – the BKS tests – using the paper-based, scheduled monthly test windows. Any awarding body that meets the on-demand criteria (see Appendix 1) may offer the tests on paper and – subject to compliance with further conditions – may also do so on screen (see Diagram 1, below).



The basic and key skills (BKS) e-assessment experience report

The test papers used in the above delivery options are designed, developed and produced by QCA. The test specification is written by QCA on behalf of QCA, CCEA and ACCAC. Each test contains 40 multiple-choice questions that are grouped in scenarios. The same bank of tests is made available to all accredited awarding bodies for national delivery. The tests are delivered, administered and marked by the awarding bodies. Some tests use predetermined pass marks, whilst others use pass marks set following the administration of the test.

Methodology

The methodology used during the project was to facilitate the work of each BKS awarding body, with the pace and rate of technical advancements dictated by each awarding body. In order to foster ongoing dialogue between QCA and the awarding bodies, a number of approaches were used, including:

- pre-implementation discussions and advice
- visits by QCA staff to centres during each pilot
- preparation of reports following each visit
- evaluation of any awarding body reports
- meetings with individual awarding bodies to discuss issues specific to their pilots
- meetings with a group of awarding bodies to discuss generic issues
- developing good practice.

In addition, as all the awarding bodies were using common tests, it was agreed that no awarding body's on-screen delivery system should generate differences in the psychometric properties of the test and any of its components. Therefore, the project team concluded that there was a need for the different awarding bodies' on-screen delivery systems to have a measure of uniformity; further, that the uniformity was to be achieved by levelling up the different delivery systems to meet the developing concept of good practice. Hence, when deficiencies in delivery systems were identified, the awarding bodies were asked to address them during the post-pilot revisions to their systems.

Use of pilot studies

The risks – to candidates, centres, awarding bodies, QCA as regulator and (possibly) to qualifications in general – that might follow the release of an inadequate or unreliable on-screen delivery system were perceived to be large. Therefore, the work was progressed through series of:

- **non-live pilots** to evaluate the individual steps in each awarding body's on-screen delivery system and the extent to which they have been joined up to produce a complete system
- **live pilots** to prove the robustness of each on-screen delivery system, awarding body support systems and the adequacy of centres' preparations for on-screen testing
- **extended live pilots** to evaluate how each awarding body's on-screen delivery system needed to be developed to deliver:
 - on-demand assessment
 - more qualifications to more centres and more candidates.

The live and extended live pilots led to the formal awards; the non-live pilots did not.

After each series of pilots, the reports of visits by QCA staff were evaluated and compared with each awarding body's own evaluation report, and discussions were held.

Feasibility studies of on-screen testing – January 2002

The purpose of the initial phase of research, known as the 'January 2002 feasibility study', was to identify the issues that would follow from the delivery of the BKS tests using IT-based systems. The products of the work were to inform the direction of future work. The key stakeholders agreed the project's scope, brief and outcomes. In particular, the 'national status' of the BKS tests was to be retained by requiring the participating awarding bodies to use the common set of test 'papers' supplied for the purpose. The outcomes were also specified. For example, there was a requirement for participating awarding bodies to provide results to candidates within 7 to 10 days of the tests having been taken. The awarding bodies were also required to work within the regulatory framework and code of practice.

All 18 awarding bodies approved to offer the BKS qualifications were invited to participate in the initial feasibility study. Participation by the awarding bodies was subject to the following conditions:

- an assessment of whether the proposed approach was practicable and that the capability, quality, methodology and feasibility criteria were satisfied
- achievement of the prime deliverable of authoring and delivering the BKS tests to centres electronically
- an assessment of whether the awarding bodies could meet QCA's deadlines
- an assessment of the previous experience of the personnel in the awarding body in managing research studies
- the cost of the research studies.

Seven awarding bodies participated in the feasibility studies. The outcomes¹ from the first phase of the feasibility studies were:

- the identification of a variety of models that might be used to deliver the level 1 and level 2 BKS tests. (The pilot also identified models that might be used to deliver the level 3 and level 4 BKS tests, but they were not developed at this point in the project.)
- the models were dependent on the volumes of candidates, frequency of testing, and the medium for delivery (on paper or on screen). A number of intermediate models (between the all-paper and all-on-screen ones) were identified
- the success or failure of the electronic delivery of the BKS level 1 and 2 tests proved to be dependent on the facilities and infrastructure available to the awarding bodies and their centres. Both were found to be significant limiting factors
- the delivery models developed by different awarding bodies varied considerably in their use of delivery technologies and the extent to which functions were made available on screen. Nevertheless, all were able to report within the 7 to 10 working days specified
- a set of general principles for on-screen assessment were developed and used to inform later work
- the project was extended to a second set of feasibility studies.

¹ See QCA's summary report on the research study project (April 2002)

Pilot studies

Non-live and live pilots – May 2002

The second series of BKS pilots involved QCA in facilitating the work of individual awarding bodies in their development of IT-based delivery systems. The main purpose of the project changed from simply demonstrating the feasibility of on-screen test delivery of BKS tests electronically to working with the awarding bodies to develop joined-up, robust on-screen delivery systems. The assessment procedures, administrative processes and technical features of the awarding bodies' on-screen delivery systems differed significantly from each other. However, they were all based on the generic model of the assessment process² and included phases for:

- design of the assessment instrument
- preparation, authoring and calibration
- entry and registration of candidates
- distribution of electronic test papers to test centres
- authentication of candidates' identity by test centres
- delivery of the electronic test to candidates by the test centres
- return of candidates' responses to the awarding bodies
- marking and processing of candidates' results by the awarding body
- data return and analysis by the awarding body
- results standardisation
- results publication and certification
- appeals.

Between May and November 2002, several of the awarding bodies undertook further series of non-live and live pilots, using their developing on-screen delivery systems and further versions of the BKS tests. The second series of pilots culminated, in November 2002, in the development of robust on-screen delivery systems by two awarding bodies.

Non-live, live and extended live pilots – November 2002 onwards

Between November 2002 and May 2003 a total of five awarding bodies undertook series of non-live, live and extended live pilots.

The two awarding bodies that had developed robust delivery systems conducted further series of live pilots and produced plans for extended pilots, which included elements of on-demand delivery. The issues arising from these series were predominantly centre-based.

Two further awarding bodies took the opportunity to introduce new delivery systems and undertook a series of live and non-live pilots. By the end of May 2003 these two awarding bodies had effectively caught up with the two leaders. Again, centre-based issues tended to dominate.

² *UIA project – final study report* (July 2002)

The fifth awarding body was able to demonstrate a prototype delivery system, but it did not comply fully with the developing good practice in on-screen assessment and was considered to be demanding of the time of centre staff.

The awarding bodies' proposals for the schedule and administration arrangements for the tests were based on the anticipated demand from candidates and centres. Hence the schedules varied across the five awarding bodies. However, a pattern emerged, in which the tests for a particular subject and level tended to take place on a common date. One awarding body allowed its centres to take the level 1 and level 2 tests on the same day, and within the test window.

An important achievement of this series of pilots was an understanding of the centre-based issues and how awarding bodies will have to work with their centres to ensure successful on-screen assessment sessions. During this period all five awarding bodies developed their own training programmes. In addition, staff from the University for Industry (Ufi), working for the Learning and Skills Council (LSC) and LearnDirect, developed a training programme and manual for the LearnDirect centres, many of which were new to the concept and practices of external assessment. The training was based on British Standard 7988. Ufi/LSC/LearnDirect required staff from LearnDirect centres to be trained by them, and part of the programme was specific to the awarding body system that they proposed to use.

With five awarding bodies undertaking on-screen pilots, and several others taking the first steps towards on-demand delivery, the supply of BKS tests became an issue. To date, the on-screen pilots have only involved the delivery of the level 1 and level 2 information technology (IT), application of number (AoN) and communication papers. The awarding bodies were required to put in place contingency plans for IT failures, plus rigorous security arrangements. Many of the awarding bodies therefore decided to schedule their on-screen tests prior to the monthly paper-based tests, so that, in the event of an IT failure, candidates could be given the opportunity to take the scheduled paper test.

The awarding bodies were provided with sets of reconfigured, cloned and unique BKS papers, plus the relevant pass mark and/or pre-determined pass mark for each test. Under the conditions of the pilots, awarding bodies were required to process results within 7 to 10 working days. Some awarding bodies provided results within 24 hours.

Extended live pilots – May 2003

Since May 2003 three awarding bodies have been undertaking 'extended' live pilots (offering the BKS tests for a period of six months or more), in which the BKS tests are available to centres on demand and on screen.

Under the 'extended' pilot arrangements, the delivery of the BKS tests to centres is at the discretion of the awarding bodies but subject to:

- maintaining the security of the test papers by having processes and procedures in place to prevent centres from retaining copies of the test papers and/or preventing centres from printing copies of the test papers

- adherence to the principles of BKS external assessment by using the national tests provided by QCA
- presenting each item/scenario and test on its IT-based delivery system, so that the test uses the same numerical order for the questions, as per the original test paper provided by QCA. Therefore, the awarding bodies may not randomise either the questions or distractors, nor make any other amendments to the questions, without prior permission from QCA
- whilst each awarding body's on-screen delivery system is different, the font type and size, use of colour, screen layout, screen navigation and so on may vary across awarding bodies, but the items/questions when presented on screen are required to replicate the paper format. For example, the scenario information must be presented with each item on screen and accompany all relevant items by presenting the on-screen paper using 'virtual paper' format
- providing pass/fail feedback to centres within 10 working days of the test being taken and providing diagnostic feedback within four weeks for candidates who fail the adult literacy and numeracy tests
- monitoring and comparing the pass rates for the on-screen papers with the paper-based delivery
- issuing certificates to successful candidates
- evaluating the non-live/live/extended live pilots by preparing reports for QCA.

By autumn 2003 five awarding bodies were either delivering BKS tests on screen or poised to do so. These extended live pilots have continued into 2004. One awarding body offers full on-demand, on-line, on-screen assessment opportunities, with immediate, profile reporting of results. The other four have developed common variations of an off-line and on-demand model, in which forthcoming assessment windows are agreed with centres and candidates are registered. Individual centres are then allowed to request an assessment session at short notice within the assessment window, when they believe that individual candidates are ready.

Independently of the awarding bodies' evaluations of their systems, QCA staff have continued to undertake monitoring visits to awarding bodies and test centres to assess whether the method of delivery offered by each awarding body is 'fit for purpose', reliable and robust. The APDP e-assessment team has managed the visits, with visits being made by staff from across the APDP, BKS and audit teams.

At the time of writing, one awarding body has been delivering between 1,000 and 1,200 on-line tests per month and another is believed to be working at a similar level.

Issues relating to the participating awarding bodies' IT-based delivery systems

- Some of the participating awarding bodies' delivery systems involve use of the internet to access the awarding body's website. Comments have been made that this might pose a security risk. Those awarding bodies that are using the internet have confirmed that they use the 'industry standard' to distribute the tests – that is, they use compressed encrypted test files and use encryption to communicate passwords and data. Some awarding bodies have said that the main perceived security risk is not from the use of the internet, but rather from fraudulent centres and/or invigilators.
- Most of the awarding bodies have reported local infrastructure issues with availability of facilities and equipment during the pilots. Initially this limited the scope of the pilots. However, recently it has been reported that as centres' confidence in the use of on-screen tests increases, the take-up by new centres has increased.
- All awarding bodies are required to have systems in place for test version control, so that centres and candidates cannot predict the content of the test. Some test centres have reported that their awarding body's procedures for test version control are inadequate. On investigation by the awarding bodies, it has been noted that the centres may have become confused, as some of the test papers look similar. However, given that the centres are unable to print copies of the papers, the centres concerned are relying on memory to make comparisons, as they are unable to make direct comparisons between test papers. The awarding bodies have reassured QCA that resit candidates have not been presented with versions of tests that they may have taken already. These awarding bodies' systems track individual candidates by the use of unique candidate IDs.
- Most of the awarding bodies provide free training to all new on-screen centres, and many have made this a mandatory component of the centre approval process. The training sessions raise general awareness of the IT-based delivery system, candidate registration, system installation, invigilation and the exam process, and provide the opportunity for hands-on experience in the use of the delivery system.
- The awarding bodies' procedures for candidate registration and authentication vary. One awarding body's procedures do not require centres to print, complete and submit candidates' registers to confirm each candidate's authenticity; the majority of the others do. However, for those systems where centres are not required to print, complete and submit candidates' registers, candidates are required to bring a photo identification card as an alternative to the register.
- Some of the participating awarding bodies' procedures require centres to undertake compliance checks on the centre's server and workstations before the test starts, whilst others do not. Those awarding bodies whose procedures do not include the need for centres to undertake compliance checks on servers and workstations have been advised that checking of equipment prior to a test session is considered to be good practice.

- All the participating awarding bodies provide advice to their test centres on the conduct of the test. Some awarding bodies' test centres have failed to follow the advice provided by the awarding body, as no exam notices were placed on the examination doors.
- The awarding bodies provide advice to their test centres on invigilation – that is, that invigilator should not be the candidate's tutor. Some awarding bodies' test centres have failed to follow this advice. A few centres have used BKS tutors to double up as invigilators, and the awarding bodies have since reinforced this good practice by reminding centres of the need to follow this policy advice.
- Each of the participating awarding bodies provides a helpline for test centres to contact in the event of the invigilator or centre manager needing assistance or advice. The awarding bodies have advised centres that access to these helplines should be easily available during the administration of the tests, by providing immediate access to a telephone in the room where the test is being conducted. Some test centres are not following this advice.
- Most of the participating awarding bodies provide advice to their test centres on the layout of the test room and workstations. During QCA observational visits, the room layout at some test centres was considered to be unsuitable for the assessment of large numbers of candidates, given the wall-to-wall layout of workstations. The awarding bodies have since advised these test centres to use BSI 7988 guidelines and have provided centres with advice on how to manage large numbers of candidates when the workstations cannot be easily rearranged and/or the availability of workstations is limited.
- Most of the awarding bodies provide advice to their test centres on the arrangements that need to be made and followed when a candidate takes more than one test on the same day. Some awarding bodies' test centres have failed to follow the advice provided by the awarding body in respect of taking more than one test on a single day, as candidates have been observed taking more than one test without appropriate rest breaks between tests.
- Many of the awarding bodies deliver the BKS tests to 'non-traditional' test centres, such as the workplace, colleges and training providers' premises. Occasionally, the workplace and training providers' premises have been considered inappropriate by the regulator, as these premises have not been designed for traditional teaching and testing purposes. However, it should be noted that candidates who took the tests found the facilities more than adequate, as did the centres themselves.
- Most of the awarding bodies' test centres do not use 'remote' testing, as many of the participating centres are schools, colleges and training organisations that have suitable facilities for candidates' use. One awarding body does use 'remote' testing arrangements in instances where candidates are unable to attend a test centre. This requires the invigilator to use a laptop to connect the candidate to the awarding body's website rather than using centre/workplace IT equipment.

- Some participating test centres have reported problems logging on and downloading items from the awarding bodies' websites. On investigation, some of the awarding bodies have found that the main source of the problem stems from unplanned changes to a centre's technical set-up. These awarding bodies have now recommended that centres must undertake a 'dummy test' after any technical set-up/hardware changes have been made, to enable issues to be highlighted and resolved prior to the delivery of the next live examination session. Other awarding bodies have found on investigation that the problems with logging on were not related to technical failings with the system, but rather a case of centres mistyping passwords. On the downloading issue, the awarding bodies have observed that the use of a large number of workstations, at certain peak times, does occasionally impact on the download time, as the system is unable to cope with the demand. These awarding bodies have advised centres to download the test/items prior to the test session and/or to use broadband connections, which alleviates the problem.
- The IT-based delivery systems vary from awarding body to awarding body. However, most use common presentational features and common functionality. The majority of candidates who have taken the tests found them to be presented clearly. However, a few candidates reported eyestrain, as the use of a small font size has caused some candidates difficulty in reading the text. Similarly, one centre visited did not have a wall clock in the room. However, the candidates are provided with an on-screen clock and are given time warnings.
- The awarding bodies have advised centres that candidates need to be given the opportunity to take navigation/practice tests. However, not all centres conform to best practice, as some candidates have not been given the opportunity to try a practice or navigation test until immediately before the actual test. These candidates have been observed as being 'confused' during the test. The awarding bodies are recommending that the opportunity to take a practice test is provided a day or so prior to the candidates' first test.
- All participating awarding bodies' systems vary, with some awarding bodies' IT-based delivery systems using the 'zoom', 'scroll' and/or 'exhibit text' features to access questions and scenarios. Participating awarding bodies have reported that this is an issue with the design of the content of the paper, as the test papers have been designed for 'pencil and paper' delivery. These papers/items do not always transfer well to on-screen testing, for instance the use of scenarios and large graphic files that cannot be easily embedded into the question. The awarding bodies have observed that only QCA is in a position to readdress this issue, by making the test content more relevant for use on screen.
- Most of the awarding bodies' systems capture candidates' responses as they proceed through the test. This is considered good practice, as candidates' responses should be retrievable in the event of technical failure. However, at present one awarding body's IT-based delivery system is unable to retrieve candidates' responses if the centre loses its connection during the test or the software unexpectedly fails. This system, however, is able to provide a pass/fail notification on the basis of how far the

candidate had got and how many correct responses the candidate submitted before the system crashed.

- All awarding bodies are required to process and report results within 10 days of the test having been taken. Some of the awarding bodies use reporting functionality that provides candidates with instant results. This functionality enables the centre manager or administrator to produce a report of all candidates' scores after the completion of the test. Other awarding bodies have decided to provide candidates with results within 24 hours, rather than immediately. This is not an issue related to the capacity of the awarding bodies' technology, but rather a policy that these awarding bodies have chosen to implement, as they scrutinise the results before release.
- All awarding bodies are required to provide adult literacy and adult numeracy candidates with detailed feedback on the test. Some awarding bodies have developed software that enables centres to provide basic skills candidates with feedback on each adult literacy and adult numeracy skills standards area; others are in the process of developing software. Those that are developing software provide candidates with written detailed feedback within four weeks of the test having been taken.
- All awarding bodies are required to have policies and procedures for those candidates requiring special arrangements. During one visit, it appeared that the awarding body's policy and procedures on special needs had caused some confusion.
- All awarding bodies are required to have contingency arrangements in place. One test centre did not have a contingency plan, that is, spare workstations for use in the event of IT failures. The centre has since been advised of the awarding body's contingency policy.
- Given the on-demand and flexible nature of testing opportunities available across the participating awarding bodies, the resit policies across the awarding bodies vary. Some test centres have reported that their awarding body has a rule that limits the number of permitted resits to three. QCA has advised the awarding bodies to make use of all the papers provided with the on-screen pilot project. Other awarding bodies have advised their test centres that they may register candidates to take the BKS tests on a weekly basis. One centre interpreted this to mean that it could re-enter candidates who fail the test on a weekly basis, until the candidate passes the test. The awarding body has since issued advice to its centres that this is not good practice and that appropriate training and support should be given to candidates between resits. Some awarding bodies have asked QCA for the time between resits to be reduced from seven days to one day; these awarding bodies were informed that this was contrary to good practice.
- All participating awarding bodies are required to undertake an evaluation exercise after each pilot stage. Some have conducted a full evaluation of the on-screen pilot. For instance, some candidates have reported a preference for the paper-based tests.

The basic and key skills (BKS) e-assessment experience report

- Most of the awarding bodies have asked for information about how and when the status of the pilot could be changed, so that on-screen assessment becomes part of the assessment regime. This is an ongoing issue and is being considered by QCA.

Summary of the pilots – January 2002 to date

A number of awarding bodies have been offering the BKS tests on screen, with three providing the BKS tests to centres on screen using on-demand models, that is, as and when required by the centres and candidates, as part of the extended live pilot arrangements. A fourth is offering the tests as part of a 'controlled' extended live pilot. The other awarding bodies are at various stages of non-live and live piloting.

The majority of the problems/issues raised during the series of live and extended live pilots were initially IT systems-based, but are now predominantly centre-based. However, there are still some issues associated with the awarding bodies' IT-based delivery systems (as indicated above). The issues raised above were 'one-offs' and are not reflective of every test session.

During the 'extended live' pilots many of the issues were addressed at centre level. Whilst some issues at centre level still need to be resolved, some of the awarding bodies are claiming that testing of the BKS tests is now quite advanced. These awarding bodies consider their IT-based delivery to be sufficiently reliable and robust to offer the BKS tests without restrictions, other than meeting the on-demand criteria (see Appendix 1).

Additional findings

To date (between September 2002 and December 2003) approximately 12,500 candidates have taken the BKS on screen with the participating awarding bodies.

The results from the on-screen pilots show that the IT, AoN and communication pass rates are considerably higher than the paper equivalent, with candidates taking the AoN and IT papers performing particularly well. The awarding bodies have reported that candidates taking the communication papers did not perform quite as well as the on-screen IT and AoN papers, though still better than the paper-based equivalent, as there was a considerable amount of text within the test papers that was not designed for on-screen delivery. Similarly, a comparison of the performance of key skills candidates against basic skills candidates in AoN/numeracy and communication/literacy has shown that the results across the two qualifications were generally similar at each level. However, the pass rates at level 2 in AoN/numeracy and communication/literacy showed that basic skills candidates perform better than key skills candidates.

As explained in the background information above, the tests used in the on-screen pilot were the same as those used in the on-demand paper-based pilot. Concern has been expressed about the comparability of the BKS tests, given that large numbers of papers are being developed. One set of papers, developed by the Basic and Key Skills Venture (BKS), was the original tests, the pass marks for which were set at post-administration standardisation meetings. This set of tests was used in the paper-based monthly scheduled test windows. The second set, developed by Goal/Education Development International plc (EDI), was the reconfigured and/or cloned versions of the unique BKS tests. The pass marks for these papers were predetermined. These papers were used in the on-screen and on-demand pilots. In

addition, it was agreed that, from January 2003, once the unique BKS tests had been taken they should be released for use in the on-screen and on-demand pilots.

Take-up of on-screen BKS tests has been increasing, with many of the awarding bodies reporting that if the status of the project could be changed, take-up would increase further. These developments indicate that the awarding bodies have taken a cautious approach in offering the BKS tests on screen. However, from the relatively small number of QCA monitoring visits, questions over the validity, reliability, accessibility and/or security of the IT-based delivery systems in use by the awarding bodies continue to be raised. It is important to note that these issues are specific to one incident and not a general issue across the pilots.

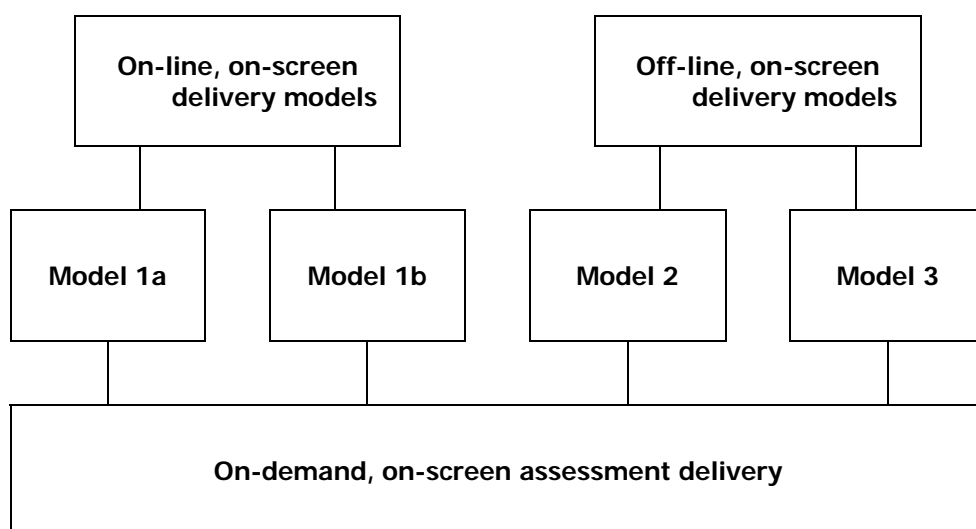
An evaluation of the key features of IT-based delivery systems used in the BKS pilots

The awarding bodies use different IT-based delivery systems, each with different features, to make the BKS tests available to centres. One system uses on-line and encrypted internet systems (models 1a and 1b in Diagram 2 below) to access the BKS tests; another uses the intranet with encrypted files/CDs (model 2) to access the tests from the local server; and a third (model 3) uses encrypted files/CDs that are free-standing and not linked to the centre's server or awarding body's server. All the IT-based delivery systems:

- deliver the BKS tests securely to test centres
- enable the awarding bodies to take greater control and offer greater flexibility to centres by varying the testing process, from registration to the reporting of results
- allow the awarding bodies to provide the BKS tests, so as to maximise the numbers of candidates who can take tests simultaneously in each centre
- allow the awarding bodies to provide the tests to centres as and when required
- enable the awarding bodies to manage the delivery of the tests so as to permit candidates taking the same subject and level, at the same time, to receive different versions of a test
- allow the awarding bodies to process and report candidates' results rapidly
- make other qualifications available to centres.

However, all the IT-based delivery systems are different and therefore have different key features and functions. As a result, each awarding body's system presents items/questions using different icons and screen formats, and requires different navigation steps to be undertaken to access the tests. This is summarised in Diagram 2.

Diagram 2: The relationship between the different delivery models



The main features of the IT-based delivery systems used to deliver the on-screen tests are listed in Table 1. An evaluation of the key features of each system is shown in Table 2. As indicated in the section above, there are still issues with each of these delivery models.

There is considered to be value in having a variety of on-screen delivery models for awarding bodies and centres to select from:

- it is possible to evaluate the strengths and weaknesses of individual features using a variety of contexts
- centres can select on-screen delivery systems to meet local needs.

Table 1: Key features of the on-screen IT-based delivery models used in the BKS tests

Model:	1a	1b	2	3
Outline description	Items are delivered to candidates on-line throughout the assessment session	Assessments are delivered to candidates on-line at the beginning of the assessment session	Assessments are made available to centres in advance and loaded onto centre's server	Assessments are made available to centre in advance on CD and loaded onto individual workstations
Steps in the delivery of assessments from the awarding body to the candidate	From awarding body's server directly to the candidate's workstation via secure internet connection	From awarding body's server directly to the candidate's workstation via secure internet connection	1. From awarding body's server to centre's server via: <ul style="list-style-type: none"> • secure internet • web access • CD 2. From centre's server to candidate's workstation	1. From awarding body to centre via removable media 2. Assessments are loaded directly onto candidate's workstation
Steps in the delivery of responses from candidate's workstation to awarding body's server	From candidate's workstation directly to awarding body's server via secure internet connection	From candidate's workstation directly to awarding body's server via secure internet connection	1. From candidate's workstation to centre's server 2. From centre's server to awarding body's server via secure internet connection	1. Centre's staff collate responses 2. From centre's server to awarding body's server via secure internet connection OR collated responses are sent to awarding body on CD
Steps in the notification of results and provision of feedback to candidates	From awarding body's server to centre's server OR to candidate's workstation via secure internet connection	From awarding body's server to centre's server OR to candidate's workstation via secure internet connection	From awarding body's server to centre's server via secure internet connection	From awarding body to centre server via secure internet connection or CD
Classification	On-screen, on-line, on-demand	On-screen, on-line, on-demand	On-screen, on-demand	On-screen, on-demand

Table 2: A comparison and evaluation of the main on-screen delivery models

Model	1a	1b	2	3
Lead time for registrations	Candidates are registered in advance for the qualification			
Notice required for registered candidates to take an on-screen assessment	No notice required	Within 24 hours	Within 24 hours	A lead time of several days is required for awarding body to send assessment and for centre to load assessment onto workstations
Speed of return of results	Potentially seconds for MCQ tests with pre-set pass marks			Potentially a few days
Form of results	All four models have the potential to report results in pass/fail, marks or grades, and profile forms			
Software installation and updates	None, other than compliance with awarding body's technical specification		Initial software is loaded by awarding body or by centre's staff. Updates are sent via internet with forthcoming assessments	All software and updates are loaded by centre's staff
Demands on centre's staff and technicians	Low (assuming adequate preplanning)		Medium (assuming adequate preplanning) Input is needed to manage distribution of assessments to workstations	High In some versions of the model, centre's staff load assessment onto each workstation, prepare 'assessment keys' and collate results for submission to awarding body
Dependence on internet	Good connections are required throughout the assessment session. Broadband is recommended	Good connections are required at the beginning and end of the assessment session. Broadband is preferred	Good connections are required for awarding body to send assessments in advance and to receive candidates' responses	Good mail or courier services required for awarding body to send assessments and receive candidates' responses
Factors limiting the quality of delivery service to candidates	Poor internet connections	Non-compliant servers and workstations	Non-compliant servers and workstations	Non-compliant workstations. Poor quality of removable media. Loss of candidates' responses
Security issues	Risk of internet hacking	Possible retention of test or item images on workstations and servers. Internet hacking	Possible retention of test or item images on workstations and servers. Potential loss of removable media in transit to or from centre	Possible retention of removable media and test or item images on workstations. Potential loss of test or removable media in transit to or from centre

Evaluation of the project methodology

The initial focus of the project was on working with awarding bodies to achieve robust delivery systems. When the first of these systems were produced, the direction of the project changed to centre-based issues. Finally, the focus moved to the management issues that on-screen assessment is producing for awarding bodies and centres.

The use of common BKS papers was taken to imply the need for a measure of uniformity in the delivery systems developed by different awarding bodies. In practice this meant that awarding bodies were encouraged to consider the developing concept of good practice in on-screen assessment.

This approach in direction was taken following QCA staff attendance at participating centres. Repeated visits were made to those centres where the assessment sessions failed, initially due to technical issues with the delivery systems, then technical and managerial failures in the centres. Each failed session raised new issues or provided new insights into previously identified issues. The successful sessions provided confirmation that earlier issues were being addressed and that the total number of issues was finite.

Some tensions arose with a number of awarding bodies when failed assessment sessions that had been observed by project staff were reported as successes or ignored altogether. Therefore the quality of some awarding body evaluation reports remains a concern. While some reports contained frank discussions of the issues encountered during the pilots, and the actions taken to address them, others tended to gloss over the issues and focus on further dissemination plans or public relations.

This approach taken with the methodology has been criticised as 'over-regulation', with the associated risk that innovation has been stifled. In contrast, it should be noted that when non-BKS qualifications and awarding bodies' own qualifications have been delivered on screen, they have not been subject to the same sequence of pilots and visits. Instead, the accreditation process involves an a priori evaluation of a description of the awarding body's system, using the BSI standard 7988. Thereafter, issues of delivery quality are dealt with through the routine (but infrequent) audits of qualifications undertaken by the regulators. It should also be noted that new accreditations of qualifications tend to be for on-paper or on-screen delivery – seldom both.

It is difficult to reconcile the criticism of 'over-regulation'. First, when comparisons are made between the developed good practice and the BSI standard 7988, it is found that there are few differences in the two lists of features that the delivery systems are expected to contain. Second, the developing delivery models are technically diverse – some have said too diverse. Third, the current project has identified new management issues that BSI 7988 did not anticipate. Lastly, the opportunity to learn from the many technical and managerial failures at the awarding body and centre levels is being lost, as far as non-BKS qualifications are concerned.

Conclusions and recommendations

The observations made during the visits to on-screen assessment sessions, especially those that failed, confirm the need for a measure of uniformity in the delivery of BKS tests on screen. Rather than the psychometric concerns that prompted this need for uniformity, concerns about the quality of the on-screen assessment experience for candidates have predominated.

The core of good practice that has been identified consists of four elements:

- desirable features of on-screen delivery systems
- advice to awarding bodies on how they should induct their centres in on-screen assessment
- advice for awarding bodies to pass on to their centres about:
 - achieving and maintaining compliance with the awarding bodies' technical specifications for on-screen delivery
 - the management of on-screen assessment sessions
 - the management of ICT resources in the centres to ensure satisfactory delivery of on-screen assessments
- identification of issues for awarding bodies to resolve as they scale up their on-screen delivery systems to accommodate more qualifications, more centres and more candidates.

These elements have been 'tested' through advice to awarding bodies considering the on-screen delivery of other qualifications.

The project has produced insights into the nature and use of e-assessment in the medium to longer terms. Links between on-screen assessment and e-marking have been identified. Models and scenarios have also been developed for the implementation of further on-screen assessment. Suggestions include:

- models for the delivery of on-screen assessments at levels 3 and above
- question types, and associated methods of marking, that will be required for the development of on-screen assessments at levels 1–5
- scenarios for the future management of on-screen assessment at the awarding body and centre levels.

Participating awarding bodies have also been able to produce the first UK comparisons of the psychometric properties of common test items that have been delivered on paper and on screen.

The value of the observations made during the visits to on-screen assessment sessions is an endorsement of the project methodology. Criticisms of 'over-regulation' are rejected.

It is recommended that the monitoring systems developed for this project are extended to all qualifications delivered using on-screen methods, and become part of the regular audit process. Audit and monitoring visits should consider the reliability, validity, 'fit for purpose' and the management principles of all on-screen delivery systems.

Awarding body issues

There is a general concern regarding the ability of all participating awarding bodies – whether they are planning to deliver BKS or other qualifications – to scale up their on-screen systems to deliver them on a national basis. **It is recommended that the practicality of awarding bodies' plans to scale up their delivery systems are kept under review.**

QCA's monitoring visits and the evaluations undertaken by the awarding bodies have indicated that candidates prefer to take their BKS tests on screen, but some have requested the choice of doing so on screen or on paper. **It is recommended that all awarding bodies are encouraged to continue to use questionnaires and interviews to assess reactions of candidates and centre staff to on-screen assessment.**

In parallel with the work reported above, at least one other project managed by QCA has raised the issue of whether there should be a national infrastructure to support e-learning and e-assessment. **It is recommended that awarding bodies developing on-screen and other e-assessment systems are included in any national initiatives or frameworks to provide IT infrastructures across different educational sectors.**

Feedback from awarding bodies and their centres has indicated the need for QCA to provide policy advice on the frequency of on-demand resits. **It is recommended that – the supply of papers permitting – awarding bodies are advised that resits should not take place within less than seven days.**

QCA staff attending some of the pilot on-screen delivery sessions have observed BKS and other qualifications being delivered in non-traditional locations as diverse as company offices and workshops, hotel bedrooms and local test centres. **It is recommended that the issues associated with the use of these 'new' assessment locations are collated, with a view to identifying and disseminating further good practice.**

QCA staff have also observed some of the special arrangements that awarding bodies and centres have made for the on-screen assessment of candidates with particular assessment requirements. So far the project staff have only been able to list reported/observed examples of need and the arrangements that the awarding bodies put in place to address them. **It is recommended that issues associated with the on-screen assessment of candidates with particular assessment requirements are reported and collated, with a view to identifying and disseminating further good practice.**

Centre-based issues

Concerns have been raised throughout this report about centre-based issues that may inhibit the take-up and management of on-screen assessment. The achievement and maintenance of technical compliance with the awarding body's specification and the management of centres' ICT resources to ensure adequate on-screen test sessions are particular concerns. **It is recommended that these issues are kept under review, with a view to identifying and disseminating further good practice.**

Issues concerning the supply of BKS test material

Feedback received from the participating awarding bodies, candidates and centre staff has indicated that the BKS test content should be designed specifically for use on screen. **It is**

recommended that consideration is given to how and when the development of on-screen BKS items and tests might be undertaken.

The need remains for large quantities of tests to resource anticipated demand for on-screen, on-demand BKS testing, and ensure adequate version control. **It is recommended that the awarding bodies are encouraged to develop and design further test items and 'item banks'. Any new items should be of similar quality to existing ones, and the bank should be capable of generating balanced BKS tests on demand for use both on paper and on screen.**

Continuing pilot status

The continuing pilot status of the BKS on-screen delivery systems remains an issue. **It is recommended that the issue of accrediting on-screen delivery systems on a non-pilot basis is considered during the next round of accreditations/re-accreditations.** The implication is that an awarding body might be granted accreditation to offer:

- the BKS qualifications on paper and on screen
- the BKS qualifications on paper, but to make the on-screen tests available as part of a controlled pilot
- the BKS tests as paper-based tests only.

Appendix 1: The on-demand criteria

External assessment of level 1 and 2 key skills, adult literacy and numeracy: advisory criteria for use of the test bank

The level 1 and 2 basic and key skills test bank provides approved basic and key skills awarding bodies with the facility to offer their centres flexible access to the tests and pass/fail feedback within 10 working days of a test being taken.

The test developer will supply awarding bodies that meet the following conditions (see 'Awarding body responsibilities prior to gaining access to the test bank', below) with a bank of test papers for use by the awarding body to meet the external assessment requirements of its centres.

The tests in the bank are not supplied with designated test dates but are intended for use by awarding bodies as and when required, on a rotational basis

The bank consists of papers that have been developed and used previously for scheduled monthly assessment windows, and new and unique papers developed for the purpose of on-demand assessment.

All tests supplied are written to the basic and key skills test specifications and are produced subject to the same set of quality assurance procedures, to ensure that a comparable level of demand is assessed from one paper to another.

The purpose of the test bank is to enable flexible delivery of the tests by awarding bodies to centres and delivery of results (pass/fail) to centres within 10 working days of the test being taken.

To support the delivery of results within this timeframe, the test developer will supply papers with recommended, predetermined pass marks.

In the case of any previously used test paper that has been the subject of an awarding meeting, the recommended pass mark will be that agreed by awarding bodies acting jointly at the awarding meeting. Once an award has been made, participating awarding bodies should add all papers that have no deleted items to the bank.

In the case of any test paper that has not been the subject of an awarding meeting, the test developer will seek a recommended pass mark from chief examiners acting for each awarding body with access to the bank at the time the paper is released into the bank. The test developer will aggregate these recommendations to determine a recommended pass mark for each paper. The test developer will monitor and keep under review the process for predetermining pass marks.

Awarding bodies' responsibilities prior to gaining access to the test bank

Before being granted access to the test bank, the awarding body must provide evidence of having developed the following systems and procedures. This evidence should include descriptions of all required systems and procedures.

The awarding body will develop systems:

- for rotated delivery of tests so that centres are not able to predict the paper their candidates will be sitting on any given assessment occasion
- to ensure that no centre receives the same paper more than once within a given period of time (this period of time will depend on the number of papers in the bank and how frequently centres use the tests)
- to ensure, as far as possible, the complete security of the papers, to ensure that centres do not retain copies and that papers are handled by centres in a secure way
- for giving pass/fail feedback to centres within 10 working days of the test being taken
- for giving unsuccessful adult literacy and numeracy candidates item level feedback on their performance. This feedback should be based on the QCA-supplied analysis of standards assessed. This feedback should be provided to centres within four weeks of the test being taken.

Once the above conditions have been met, the test developer will supply:

- a bank of papers for level 1 and level 2 adult literacy/key skills communication; adult numeracy/key skills application of number and key skills information technology
- recommended marks for each paper
- for each adult literacy and numeracy paper, analysis of the standard assessed by each item, to enable centres to give formative feedback to unsuccessful candidates.

While offering tests from the bank, the awarding body will:

- **not** reveal to centres the number of tests in the bank
- make use of all papers in the bank
- when requested by the test developer, retire papers from the bank
- add each paper from the scheduled monthly test windows into the bank once a pass mark has been determined, provided that no item deletions have been made
- issue instructions to centres that they should not retain copies of the papers and ensure that, after use, all papers are returned to the awarding body or are securely destroyed by the centre
- for adult literacy and numeracy candidates only, certificate successful candidates within four weeks
- monitor and evaluate the on-demand method for delivering tests and provide statistics and reports for the test developer/regulator as requested.

On-screen delivery

Awarding bodies are encouraged to develop their own systems for delivery of the bank of tests on screen. On-screen delivery is subject to initial QCA evaluation of the reliability and security of the awarding body's systems.