

# **SOLICITING FEEDBACK FOR A BRITISH STANDARD CODE OF PRACTICE FOR USING I.T. IN DELIVERY OF ASSESSMENTS**

**John Kleeman, David Keech and  
Stephen Wright**

# **Soliciting feedback for a British Standard Code of Practice for Using IT in Delivering Assessments**

John Kleeman  
Question Mark Computing Ltd  
john.kleeman@qmark.co.uk

David Keech  
British Standards Institution  
david.keech@bsi.org.uk

Stephen Wright  
Edexcel Foundation  
stephen.wright@edexcel.org.uk

all on behalf of BSI panel IST/43/-/1

## **Abstract**

The British Standards Institution has set up a panel, IST/43/-/1, to draw at a new British Standard Code of Practice for the use of IT in delivering assessments (BS 7988). This paper describes the work in progress and plans for creating this standard.

## **Introduction to the Standard**

The British Standards Institution have set up a formal committee, IST/43, to work in the area of learning technologies standards. IST/43 has members from government, business and education, and has two main roles. One role is to contribute to international ISO and IEC standards work, via the ISO/IEC JTC 1/SC36 committee. And the other role is to consider and work on any British standards work.

As part of this second role, IST/43 have set up a panel (subcommittee) called IST/43/-/1 to produce a specific standard in the area of computer assisted assessment. This standard is to be called: BS 7988 - A Code of Practice for the use of Information Technology in the Delivery of Assessments. John Kleeman, of Question Mark, is the convenor (chairman) of the panel.

The following organisations have been part of the initial development of the plan for the standard

- CETIS / Sheffield Hallam University
- Department for Education and Employment
- Edexcel Foundation

- Guildford Educational Services
- Intertrust Technologies International
- NCS Pearson UK
- Question Mark Computing Limited
- Qualifications and Curriculum Authority (QCA)
- Simulacra Media
- Sun Microsystems
- Sunderland University
- TÜV Product Services
- UK CAA (Computer Assisted Assessment) Centre, Luton University
- University of Cambridge Local Examination Syndicate (UCLES)

Several other organisations are also planning to become involved at the time of writing this paper.

The aim of the standard is to describe good practice in using computers and other information technology in delivering assessments. Our formal, approved scope is provided as part two of this paper. The standard does not cover authoring assessments, simply their safe and effective delivery.

At present, each organisation that uses computers for exams sets their own guidelines and works out their own practices. This is mostly because use of IT has exploded so rapidly that different sectors have not had time to communicate with each other. This new standard will set some guidelines taking the best parts of existing practices, and enabling users of computer exams to feel confident that they are fair and secure. It will build on the good work that has been done by many pioneers in computer assisted assessment, and set practical guidelines that organisations can use in the future.

The standard is currently in its drafting period. After a competitive tender, BSI has appointed Guildford Educational Services to work with the panel to produce a draft of the standard by September 2001. Included within the drafting process is consulting widely and referring to prior work in this area.

Assuming all goes to plan, the draft of the standard will be issued for formal public comment by BSI in October; and if there are no substantial objections, it's hoped that the final standard will be published around April 2002.

Although the initial work is entirely in the context of a British Standard, it is possible that the work could continue to extend the standard to apply into Europe and internationally.

The panel appreciates that many of the UK and world experts in computer assisted assessment will be at the conference, and would welcome all input, either at the conference or by e-mail to the authors as above. Particularly of interest would be knowledge of any existing documents in the area; the panel is aware of many of these, but may not be aware of all. Also if any organisations or institutions have internal documents on how they deliver computer exams that they would be willing to make available to the panel, these would be of great interest.

## Our Approved Scope

*(The following is abstracted from the approved scope for BS7988, proposed by IST/43/-/1 and approved by the BSI.)*

### **Who the standard is aimed at**

The main envisaged users of the standard are organisations that are delivering tests, assessments or exams by computer, for example:

- Universities, colleges, learning centres, exam centres and schools who assess their students, both for diagnostic and summative (final exam) purposes;
- Exam boards who deliver GCSE and other school exams, and those who deliver exams in vocational subjects;
- Producers and deliverers of IT certification tests;
- Training companies, large employers and governmental/military organisations providing tests as part of vocational education and training;
- Organisations providing tests required for regulatory purposes, for example a financial services company might be required to test its sales people for product knowledge, or a pharmaceutical company might be required to test its staff on safety procedures;
- Open learning and distance learning material providers including online universities and commercial publishers and distributors of learning materials;
- Producers of question banks, quizzes and 'tests' available as revision aids.

### **Other interested stakeholders in the standard include:**

- Test takers (who want fair tests);
- Suppliers of IT for assessment;
- Teachers and trainers;
- Outsiders who rely on the results of tests, e.g. employers, education and training consultants, regulatory authorities (e.g. QCA, ACCAC, SQA and QAA), external auditors, funding organisations and government.

The standard is a “British” standard, and needs to take into account issues relevant to Wales, Scotland, Northern Ireland and England. The standard only applies to UK testing, but takes into account international factors (including any relevant European legislation), so that it can in due course be considered for international adoption. Internationally acceptable language is used (e.g. “score” not “mark”).

### **Parts and conformance**

The standard is divided into parts, aimed at different audiences. One part is aimed at exam distributors (organizations responsible for delivering exams), and another part is aimed at exam centres (places where exams are taken).

Both parts are codes of practice, and include conformance criteria, with conformity assessment being achieved by documenting and then auditing that the defined procedures are followed.

The primary conformance criteria relate to “high stakes” testing, where the results of the assessment can have important implications for people or organisations – for

example someone's job or a major qualification can depend on it. The criteria also give guidelines on the procedures that are appropriate for lower stakes testing (e.g. self-assessments or assessments where the results are less critical).

The aim of the standard is to set out principles and good practice but not to specify technical requirements in detail. For example, it is possible to meet the standards using Windows PCs or UNIX or other computers. And file formats are not included in the standard. An organisation complies with the standard if it follows the correct procedures, but can use a variety of technological approaches to get there. The standard covers things that technology needs to do, but not precisely how it does it.

### Assessment lifecycle

Although assessment procedures vary, the typical life cycle of assessment consists of some or all of these steps:

- (a) Design of outcomes/assessment methodology
- (b) Preparation/Calibration
- (c) Pre-Registration (includes payment)
- (d) Distribution**
- (e) Authentication (includes identification)**
- (f) Delivery**
- (g) Response return**
- (h) Instant Scoring and/or Feedback**
- (i) Data Return**
- (j) Analysis
- (k) Appeals
- (l) Certification

The standard considers stages (d) through (i) but does not consider other stages of the life cycle. With regard to scoring of assessments, if this is done immediately at the time of assessment, it is covered. If it is done at a later time, e.g. by manual grading of essays after the event, it is not covered.

#### **For the avoidance of doubt, the standard does cover:**

- Questions that can be scored by computer (e.g. multiple choice) and questions that cannot normally be scored by computers (e.g. essays).
- Tests that are used for diagnostics, self-assessment and remediation and which include feedback, as well as straight exams.
- Tests used in education (school and university), training, compliance and guidance.

#### **For the avoidance of doubt, the standard does not cover:**

- Pedagogy. How assessment should be done is outside the scope; how it should be implemented on a computer is within the scope.
- Test and question design. It is obviously crucially important that tests and questions are written well, however they are delivered. But this is outside the scope of the standard.

- Scoring. The standard does not describe how tests should be scored or accumulated or weighted, but should limit itself to confirming that any scoring algorithm desired is implemented correctly.
- Communication of results. The standard does not cover how the results of tests should be communicated with test takers, nor how they are stored in the long term.
- Results analysis. The standard does not cover how the answers and results should be analysed or reported on, or which statistical measures should be used.
- Tests used for recruitment and psychological aptitude tests.

The standard covers assessments within the mainstream of computer based assessment, and does not cover use of specialist technology which is only of interest in particular markets (for example flight simulators).

## Detailed list of issues

The standard sets out the procedures and practices that exam distributors and exam centres should follow when delivering computerized assessments. Although there is some commonality, in general, issues that apply to both paper and computer assessments are not dealt with; the aim of the standard is to mostly cover the new issues that apply in computerizing assessments. The following main issues are covered, though this list is not exhaustive.

## Some prerequisites and themes apply across the life cycle

- Test takers should have practice on the type of computers, software and questions used in advance.
- Security.
- Data protection.
- Accessibility and fitness for target group.
- Audit trails.
- Fallback procedures are needed for all eventualities.

### **Issues that apply particularly in Distribution:**

- Security issues need to be considered to protect against early or unauthorized access.
- Confirmation that the test is delivered/distributed electronically “intact”.

### **Issues that apply particularly in Authentication:**

- How the test taker is identified and that identification attached to their answers.
- How it's determined that the test taker is who they say they are.
- It also needs to be considered how the content of the exam can be checked to be what it is supposed to be.
- Authentication of invigilator/proctor (and any other actor).

### **Issues that apply particularly in Delivery:**

- The software needs to be proof against user errors – for example it should not be possible accidentally to select two answers in a single answer multiple choice question.
- Equipment must be in working order with equal speed of delivery.
- Opportunity for familiarization.
- Instructions on how to answer.
- Quality of user interface.
- Security issues to protect against early or unauthorized access, and to confirm that it's the right test.
- Issues relating to distance apart of computers and people seeing each other's screens, and also protection against unauthorized communication via computer.
- Some aspects of randomisation.
- Measures may need to be taken to prevent knowledge of the questions being copied by the test taker to others.
- The availability or not of aids such as spell checkers, calculators, open internet access and other resources should be defined by the test delivery organisations (and not permitted unless allowed).
- If any time limit is set it must be enforced accurately and securely.
- The confidentiality of the right answers.
- Health safety and environment - effect of computer, noise etc.
- Accessibility issues for people with special needs.
- Procedures for supervision, breaks in tests and handling of incidents such as computer failures during the examination.
- Robustness of software and the "locking down" of the computer desktop as a security issue.

#### **Issues that apply particularly in Response return:**

- Access control of responses.
- It must be possible to show an audit trail of how one person's answers become stored, so that if someone appeals against their results, one can know their questions and answers, and not only their overall score.
- If data is transferred at a distance, the integrity and security of this needs confirming.
- Backup of data in case of failure of computer systems.

#### **Issues that apply particularly in Instant scoring and feedback:**

- The confidentiality of the right answers.
- Some way of checking/auditing computer (or manual) scoring may be necessary.
- If the computer assigns a grade or pass/fail result, the procedure for this and how it is documented and checked should be covered.

#### **Issues that apply particularly to Data return:**

- The precise format of data export is not in scope, but the principles of data return are, including integrity, fairness and security.