

Code of practice for the use of information technology (IT) in the delivery of assessments

ICS 03.120.20; 35.240.99

Committees responsible for this British Standard

The preparation of this British Standard was entrusted to Technical Committee IST/43, Information technology for learning, education and training, upon which the following bodies were represented:

- Aslib (The Association for Information Management)
- British Educational Communications and Technology Agency
- BTexact Technologies
- City and Guilds
- Edexcel Foundation
- Education Online
- eLoki
- FD Learning
- IBM United Kingdom
- Inter Trust Technologies International
- NCS Pearson
- Open University
- Qualifications and Curriculum Authority
- Question Mark Computing Limited
- Qwiz Inc.
- RWCS
- Scottish University for Industry
- Sheffield Hallam University on behalf of the Centre for Educational Technology Interoperability Standards
- Simulacra Media Limited
- Sun Microsystems
- TekniCAL Limited
- Telemetry Associates
- TÜV Product Services
- Ufi Limited
- University of Cambridge Local Examinations Syndicate
- University of Sunderland
- University of Teeside
- Wide Learning Limited

The following organizations were not represented on IST/43 but contributed to the development of this standard:

- Centre for Interactive Assessment Development, University of Derby
- Computer Assisted Assessment Centre, University of Luton
- Department for Education and Skills
- Guildford Educational Services Limited
- Northern Ireland Council for the Curriculum, Examinations and Assessment
- Qualifications, Curriculum and Assessment Authority for Wales
- Scottish Qualifications Atuhority
- The Chalkface Project Limited
- University of Strathclyde
- XiSEC Information Security Management Services

This British Standard, having been prepared under the direction of the DISC Board, was published under the authority of the Standards Policy and Strategy Committee on 28 March 2002

© BSI 28 March 2002

Amendments issued since publication

Amd. No.	Date	Comments

The following references relate to the work on this British Standard:
Committee reference IST/43
Draft for comment 01/641930 DC

ISBN 0 580 39222 8

Contents

	Page
Committees responsible	Inside front cover
Foreword	ii
1 Scope	2
2 Normative references	4
3 Terms and definitions	4
4 Guiding principles	6
5 Interface between assessment content and IT delivery	7
6 IT delivery of assessments — general	8
7 Software for IT delivery of assessments — navigation and usability issues	12
8 Instant automated scoring, result determination and feedback	14
9 Preparation and transmission of assessment content and correct responses	16
10 Software and procedures for recording and transmission of candidates' details and responses and for their storage by assessment distributors	17
11 Provision of instructions and assessment-specific information for assessment centres	19
12 Provision of information and practice material for candidates	20
13 Equipment and facilities at assessment centre, including storage of responses	21
14 Staffing of assessment centre	23
15 Preparation for the assessment session	25
16 Conduct of assessment session	26
17 Procedures relating to emergencies, technical failures and irregularities	27
Annex A (informative) Scenarios	28
Bibliography	30

Foreword

This code of practice has been prepared by Technical Committee IST/43, Information technology for learning, education and training which comprises people with a wide experience of IT-delivered assessments, from a variety of different companies and organizations.

Growth in the power and capabilities of information technology (IT) has led to the increasing use of IT to deliver, score and record responses of tests and assessments in a wide range of educational and other contexts. Suitably used, IT delivery offers advantages of speed and efficiency, better feedback and improvements in validity and reliability, but its increased use has raised issues about the security and fairness of IT-delivered assessments, as well as resulting in a wide range of different practices.

The aims of this British Standard *Code of practice for the use of information technology (IT) for the delivery of assessments* are to provide a means of:

- showing that the delivery and scoring of the assessment are fair and do not disadvantage some groups of candidates, for example those who are not IT literate;
- showing that a summative assessment has been conducted under secure conditions and is the authentic work of the candidate;
- showing that the validity of the assessment is not compromised by IT delivery;
- providing evidence of the security of the assessment, which can be presented to regulatory and funding organizations (including regulatory bodies in education and training, in industry or in financial services);
- establishing a consistent approach to the regulations for delivery, which should be of benefit to assessment centres who deal with more than one assessment distributor;
- giving an assurance of quality to purchasers of “off-the-shelf” assessment software.

It is envisaged that the achievement of these aims will enhance the status of IT-delivered assessments and assessment software and encourage their wider use in situations where they are beneficial.

Users’ attention is drawn to any existing legislation covering the subject matter of this British Standard. The following British Acts of Parliament are relevant:

The Disability Discrimination Act 1995 [1];

The Special Educational Needs and Disability Act 2001 [2],

The Data Protection Act 1988 [3];

The Welsh Language Act 1993 [4];

The Health and Safety (Display Screen Equipment) Regulations 1992 [5];

The Health and Safety at Work etc. Act 1974 [6];

The Freedom of Information Act 2000 [7].

As a code of practice, this British Standard takes the form of guidance and recommendations. It should not be quoted as if it were a specification, and particular care should be taken to ensure that claims of compliance are not misleading.

A British Standard does not purport to include all the necessary provisions of a contract. Users of British Standards are responsible for their correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Summary of pages

This document comprises a front cover, an inside front cover, pages i and ii, pages 1 to 31 and a back cover.

The BSI copyright notice displayed in this document indicates when the document was last issued.

Introduction

This British Standard code of practice is intended for organizations involved in the use of IT for the delivery of assessments, including (but not limited to):

- universities, colleges, learning centres, and schools who assess their students, for diagnostic, formative and summative (final test) purposes;
- awarding bodies who deliver GCSE and other educational examinations;
- awarding bodies, professional bodies, industry organizations and others who deliver examinations or assessments in vocational subjects;
- producers and distributors of IT certification tests;
- assessment centres which administer assessments locally;
- open learning centres and other organizations which may provide assessment facilities on an occasional basis;
- training companies, employers and governmental/military organizations providing assessments as part of vocational education and training;
- organizations providing assessments required for regulatory purposes, for example a financial services company might be required to assess its sales people for product knowledge, or a pharmaceutical company might be required to assess its staff on safety procedures;
- open learning and distance learning material providers, including on-line universities and commercial publishers and distributors of learning materials, who incorporate assessments in their material;
- producers of item banks, quizzes and “tests” available as revision aids.

This British Standard is also relevant to developers of software for delivering assessments, who will need to provide software which enables their clients to comply with this British Standard.

This British Standard distinguishes three main roles in the IT delivery of assessments:

- assessment sponsors, responsible for assessment content and award of certificates;
- assessment distributors, responsible for delivering assessments via IT including developing or specifying the delivery software;
- assessment centres, where the assessments are taken.

The roles involved in the different stages of the assessment life cycle (see **1.3**) are combined in different ways by the organizations involved and not all organizations are involved in all roles. Thus, a university may combine the roles of assessment sponsor, assessment distributor and assessment centre, whilst the assessments distributed by an awarding body (as assessment sponsor and distributor) are normally taken in assessment centres which are separate organizations. This British Standard therefore recommends the action to be taken in relation to each of these three roles, irrespective of the type of organization undertaking the role. The scenarios in Annex A give examples of how this can work in practice.

The roles of the different organizations are often inter-related, especially in high-stakes assessment. For example, the assessment distributor or awarding body might have an overall responsibility for the integrity of the whole process, including monitoring assessment centres. Assessment centres need to comply with the general regulations of assessment distributors, as well as with those which are specific to the use of IT.

It should be noted that some of the organizations to which this British Standard is relevant will also need to comply with the requirements of regulatory authorities in relation to the design and conduct of assessments, including pedagogical aspects. Where the candidates are employees of the assessment centre, there are also more specific legislative requirements.

1 Scope

1.1 General

This British Standard gives recommendations on the use of information technology (IT) to deliver assessments to candidates and to record and score their responses. Its scope is defined in terms of three dimensions: the types of assessment to which it applies, the stages of the assessment “life cycle” to which it applies and this British Standard’s focus on specifically IT aspects.

1.2 Types of assessment

This British Standard is relevant to a wide range of assessments, including:

- assessments used in education (both compulsory and post-compulsory), training and compliance (e.g. compliance with legislation relating to health and safety or financial services);
- assessments of knowledge, understanding and skills (i.e. “achievement tests”), but not psychological tests of aptitude and personality;
- high-stakes assessments and examinations and also low-stakes assessments used for feedback on progress, identification of learning needs, self-assessment and remediation;
- assessments which include feedback, as well as those which provide only a result;
- both fixed-date and on-demand assessments;
- both items which can be scored by computer and the delivery (but not scoring) of items or tasks which are delivered using IT, but have to be referred to a human marker for scoring; however, the use of IT for scoring lengthy free-text responses (“essays” and similar) is excluded;
- a wide range of computer-scorable assessments, including not only “objective tests” (multiple-choice and other item types), but also assessments of keyboarding skills and software use.

NOTE Items can include graphics, multimedia and access to data, case studies etc.

This British Standard does not, however, cover the use of specialist simulations (e.g. of industrial processes or driving/piloting).

- assessments taken in universities, colleges, schools and assessment centres and also assessments taken in less formal settings, including learning centres and in the workplace;
- assessments which are set, taken and scored within a single organization (for example a university or a company) and assessments taken in an organization separate from the one responsible for the assessment content;
- assessments delivered in a variety of ways, including on-line, on local networks and on stand-alone computers.

Clause 9 of this British Standard is also applicable to the use of IT for the transmission of candidates’ coursework.

1.3 Assessment life cycle

Although assessment procedures vary, the typical life cycle of assessment consists of the following steps:

- a) identification of need to assess;
- b) design of outcomes/assessment methodology;
- c) preparation and calibration;
- d) pre-registration (includes payment);
- e) distribution¹⁾;
- f) authentication (includes identification)¹⁾;
- g) delivery¹⁾;
- h) response return¹⁾;
- i) scoring, result determination and/or feedback¹⁾;
- j) data return¹⁾;

¹⁾ Covered by this British Standard.

- k) analysis;
- l) appeals;
- m) certification.

Preparation of assessment content is outside the scope of this British Standard, but where pretesting is undertaken using IT for delivery, the relevant clauses of this British Standard should be taken into consideration.

1.4 Focus on specifically IT aspects

The emphasis throughout this British Standard is on the additional or different measures which should be taken as a result of the introduction of IT. Measures which are common to both paper and IT-delivered systems are either omitted or covered in only general terms.

This British Standard does not cover the purely pedagogical aspects of assessment preparation or scoring, but does cover the use of IT to implement pedagogical decisions and the interface between assessment content and IT delivery. The quality of the assessment content and the scoring procedures are of crucial importance to the validity and reliability of the final result, but fall outside the scope of this Standard. Compliance with this British Standard does not indicate that the assessment is pedagogically sound. This British Standard should not be interpreted as reducing the need for assessment sponsors to follow good practice in developing assessments, including standardization and attention to validity and reliability.

This British Standard does not cover scoring by human markers, but does cover:

- IT scoring which may be subject to later review by human markers;
- the IT transmission and delivery of assessments partly or wholly to be scored by human markers and the return of the resulting responses for scoring.

This British Standard covers the use of IT to provide automatic feedback and instant results, but not result-determination requiring human decision-making or intervention.

1.5 Focus on principles

The aim of this British Standard is to set out principles and good practice, but not the details of the means by which they are to be achieved. It is therefore possible to follow the recommendations using a variety of technological or procedural approaches. This British Standard is not specific to any particular hardware or software platform.

In many areas the principles outlined in this British Standard will be supplemented by the specific regulations of assessment distributors.

1.6 Compliance

Assessment sponsors, assessment distributors and assessment centres may claim compliance with this British Standard if they comply with all the clauses or subclauses applicable to their role (see table below). Notes to the clauses indicate the role(s) to which each clause or subclause is applicable.

This British Standard is applicable to both high-stakes and low-stakes assessments, but some clauses or subclauses are applicable only to high-stakes assessments; this is indicated in Table 1.

Table 1 — Application of BS 7988

Role	Assessment type	Relevant clauses or sub-clauses
Assessment sponsors	High-stakes and low-stakes	5.1
Assessment distributors	High-stakes	5.2, 6 to 12 inclusive
Assessment distributors	Low-stakes	5.2, 6.1, 6.2, 6.3, 6.5, 7.1, 7.2, 7.3, 7.4, 7.5, 8.1, 8.2, 8.3, 8.4, 9.1, 9.2.1, 10.1, 10.2, 11.1, 11.2, 12.1, 12.2 and 12.3.
Assessment centres	High-stakes	13, 14, 15, 16 and 17 inclusive
Assessment centres	Low-stakes	13.1, 13.2, 13.3 and 13.4

The scenarios given in Annex A illustrate how different types of organization might need to comply with different clauses of this British Standard.

2 Normative references

The following normative documents contain provision which, through reference in this text, constitute provisions of this British Standard. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the publication referred to applies.

BS ISO/IEC 17799:2000, *Information technology — Code of practice for information security management*.

BS EN ISO 9241-3, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 3: Visual display requirements*.

BS EN ISO 9241-4, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 4: Keyboard requirements*.

BS EN ISO 9241-5, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 5: Workstation layout and postural requirements*.

BS EN ISO 9241-9, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 9: Requirements for non-keyboard input devices*.

3 Terms and definitions

For the purposes of this British Standard the following terms and definitions apply.

3.1

assessment

test, examination or similar, designed to assess a candidate's knowledge, understanding or skills in a defined area

NOTE Can be high-stakes or low-stakes (see 3.11 and 3.17).

3.2

assessment centre

formally designated location where assessments are taken, not necessarily used exclusively for assessments

NOTE An assessment centre has staff with responsibility for the conduct of assessments and the security of material. Assessment centres are often, but not always, remote from the assessment distributor and part of a different organization. In high-stakes assessment, centres are normally subject to the approval and monitoring of the assessment distributor. Some assessments, especially those which are low stakes, can be taken outside assessment centres.

3.3

assessment distributor

organization responsible for delivering assessments (especially high-stakes assessments) via information technology, including overseeing assessment centres, but not necessarily responsible for assessment content

3.4

assessment form

alternative version of an assessment; two or more different assessment forms each contain a different mix of items but are designed to assess the same knowledge, understanding or skills

3.5

assessment software

software which presents the assessment to the candidate and records the candidate's responses

NOTE Can be generic (capable of running numerous different assessments) or integral with assessment content. Usually (but not necessarily) separate from software to develop and store the assessment items, log candidate entries, manage assessment sessions etc.

3.6

assessment sponsor

organization responsible for determining assessment content and awarding certificates

NOTE Assessment sponsors often delegate assessment delivery to an assessment distributor.

3.7**assistive technology**

devices which facilitate the use of IT by people with disabilities, especially sensory or physical impairment

3.8**candidate**

anyone taking an assessment, whether high- or low-stakes

3.9**correct response**

predetermined response which will attract a positive score

NOTE Use of this term is not intended to imply that there is only one possible correct response per item. See also 3.18 and 3.20.

3.10**delivery platform**

computer hardware, system software and, where applicable, telecommunication software and associated features which support the delivery of assessments to the candidate and the return of computer held information (e.g. results) to a specific location for scoring purposes

3.11**high-stakes assessment**

assessment where the result is of importance to the candidate or to an organization and therefore has to be reliable and valid

NOTE High-stakes assessments include most end-of-course educational examinations leading to a certificate (and also module assessments which contribute to an overall result), proficiency and licensing examinations. Assessments can also be considered high-stakes if the evaluation of an organization depends on the performance of its employees or students, as might be the case with safety assessments. See also 3.17.

3.12**ICT**

information and communications technology

3.13**IT**

information technology

3.14**invigilation**

process of overseeing the conduct of an assessment, especially to ensure that there is assessment security and no cheating

NOTE Proctoring is an alternative term for invigilation.

3.15**invigilator**

person responsible for invigilation

NOTE Proctor is an alternative term for invigilator.

3.16**item**

smallest separately identified question or task within an assessment

NOTE Can be a closed-response question, an open-ended question, an essay or the performance of a task (e.g. “centre a heading in a given passage of text”).

3.17**low-stakes assessment**

assessment where the result is of lower importance to the candidate or to a relevant organization

NOTE Formative and diagnostic assessments are usually low-stakes. See also 3.11.

3.18

open-ended response

response which is not limited to choosing from a list of options

NOTE Open-ended response items are most likely to require text input, but can require numerical input or computer graphics. See also 3.9 and 3.20.

3.19

reliability

consistency with which an assessment measures

EXAMPLE An assessment will have low reliability if two assessment forms are of unequal difficulty or coverage or if there are errors in the scoring procedures or in the reporting of scores.

3.20

response

candidate's answer to an item

NOTE Responses can consist of the selection of one or more from a given set of options, open-ended IT-scorable responses, more complex IT-scorable activities or production of text which has to be marked by a human, but can be transmitted by IT. See also 3.9 and 3.18.

3.21

result

outcome of the assessment taken by the candidate

NOTE A result can be expressed as pass/fail, as a grade, as a percentage or in other ways.

3.22

score

numerical measure of the candidate's achievement in the assessment as a whole or on the individual items

3.23

validity

extent to which an assessment achieves its aim by measuring what it is supposed to measure and producing results which can be used for their intended purpose

NOTE An assessment has low validity if the results are unduly influenced by skills which are irrelevant to the stated aims of the assessment.

3.24

workstation

equipment provided for an individual candidate at an assessment centre, including IT hardware, seating and desk or table space

4 Guiding principles

The guiding principles which underpin this code of practice and are relevant to all the organizations concerned are as follows:

- a) delivery or scoring of assessments using IT should not result in any reduction in the assessment validity or reliability;
- b) computer delivery and scoring should be fair to all candidates and, as far as possible, should not disadvantage any candidate as a result of factors which are irrelevant to the knowledge, understanding or skills being assessed (such factors include lack of advanced computer skills, if these are irrelevant);
- c) where disabled candidates are involved, reasonable adjustments should be made for their needs, so that no candidate is placed at substantial disadvantage;

NOTE Attention is drawn to The Disability Discrimination Act 1995 [1] and The Special Educational Needs and Disability Act 2001 [2].

- d) the health and safety of candidates (or of those developing, arranging or conducting the assessments) should not be put at risk;
- e) where the confidentiality of assessment content and correct responses is an important requirement, there should be procedures to maintain it throughout the assessment cycle; this should also extend to the trialling of items;

f) the confidentiality and integrity of candidate data should be maintained throughout;

NOTE Attention is drawn to the Data Protection Act 1998 [3].

g) an audit trail should be maintained, so that any queries or irregularities can be investigated;

h) back-up facilities and fall-back procedures should be in place to minimize disruptions so that the candidate is not disadvantaged, especially for high-stakes assessment.

5 Interface between assessment content and IT delivery

NOTE This clause concerns the inter-relationship between assessment content and assessment software and is relevant both to assessment sponsors, who define the pedagogical requirements of the assessment, and to assessment distributors, who are responsible for developing or customizing the software. Assessment sponsors should comply with 5.1 and assessment distributors with 5.2.

This clause is applicable to both high-stakes and low-stakes assessments and to the use of both generic assessment software (capable of running a range of different assessments) and software which is specific to a single assessment or group of assessments.

5.1 Responsibilities of assessment sponsors

5.1.1 Assessment sponsors should ensure that they have sufficient familiarity with the software which is being considered or which is to be used and the associated delivery platform to be able to:

- a) understand their advantages and limitations;
- b) appreciate the likely effect of IT delivery on the validity and reliability of the assessment and of individual items;
- c) use software features (e.g. item types, multimedia elements) relevant to the intended assessment;
- d) identify assessments for which IT delivery is unsuitable or should be supplemented by other assessment methods.

5.1.2 Assessment sponsors should specify clearly the parameters required for each assessment, including:

- a) number and type of items to be used;
- b) how the items are selected for each assessment session (e.g. fixed assessment form, computer selection from a bank, any constraints on selection);
- c) any time limit;
- d) any restraints on navigation between items (see 7.1);
- e) assessment regulations, including permitted and prohibited resources (see 6.4.3 and 7.3.3);
- f) scoring rules for individual items and for the calculation of the overall result (if applicable), including rules for the scoring of open-ended items (see 8.1 and 8.2);
- g) feedback to be provided (see 8.3).

5.1.3 Assessment sponsors should try out the assessment (i.e. the combination of software and content) during development and again before operational use to verify that all aspects of the delivery, scoring and feedback operate as intended and in accordance with pedagogical requirements.

5.1.4 In developing assessment content, assessment sponsors should consider issues relating to candidates with disabilities, including:

- a) whether a non-IT alternative should be provided for such candidates;
- b) the effect of using assistive technology on the validity of the items (for example where the wording of a text alternative to a graphic changes the nature of the item).

5.2 Responsibilities of assessment distributors

Assessment distributors should provide assessment sponsors with:

- a) full information about the capabilities, limitations and features of the intended software and the associated delivery platform that is relevant to the pedagogical aspects of the assessment;
- b) a checklist of the parameters to be specified by the assessment sponsor (see 5.1.2).