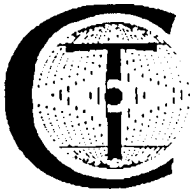

INTERNATIONAL TEST COMMISSION
COMMISSION INTERNATIONALE DES TESTS

International Guidelines for test-use

Version 2000

The Council of the International Test Commission formally adopted the Guidelines at its June 1999 meeting in Graz, Austria.
The European Federation of Professional Psychologists Associations' Task Force on Tests and Testing also endorsed the Guidelines at its July 1999 meeting in Rome.

Contents

ACKNOWLEDGEMENTS

The Guidelines were prepared for the ITC Council by Professor Dave Bartram. The author is grateful for the assistance provided by Iain Coyne in the execution of this project and is grateful to the following individuals who took part in the 1997 Dublin workshop and who provided such valuable input to the development of the present Guidelines.

Ms Dusica Boben, Produktivnost, SLOVENIA;
Mr Eugene Burke, British Psychological Society, ENGLAND;
Dr Wayne Camara, The College Board, USA;
Mr Jean-Louis Chabot, ANOP, FRANCE;
Mr Iain Coyne, University of Hull, ENGLAND;
Dr Riet Dekker, Swets and Zeitlinger, NETHERLANDS;
Dr Lorraine Eyde, US Office of Personnel Management, USA;
Prof Rocio Fernandez-Ballesteros, EAPA, SPAIN;
Mr Ian Florance, NFER-NELSON, ENGLAND;
Prof Cheryl Foxcroft, Test Commission of South Africa, SOUTH AFRICA;
Dr John Fremer, The College Board, USA;
Ms Kathia Glabeke, Commissie Psychodiagnostiek, BELGIUM;
Prof Ron Hambleton, University of Massachusetts at Amherst, USA;
Dr Karin Havenga, Test Commission of South Africa, SOUTH AFRICA;
Dr Jurgen Hogrefe, Hogrefe & Huber Verlagsgruppe, GERMANY;
Mr Ralf Horn, Swets and Zeitlinger, GERMANY;
Mr Leif Ter Laak, Saville and Holdsworth Ltd, ENGLAND;
Dr Pat Lindley, British Psychological Society, ENGLAND;
Mr Reginald Lombard, Test Commission of South Africa, SOUTH AFRICA;
Prof Jose Muniz, Spanish Psychological Association, SPAIN;
Ms Gill Nyfield, Saville & Holdsworth Ltd, ENGLAND;
Dr Torleiv Odland, Norsk Psykologforening, NORWAY;
Ms Berit Sander, Danish Psychologists' Association,DENMARK;
Prof Francois Stoll, Federation Suisse des Psychologues, SWITZERLAND.

The author is also grateful to the many other individuals and organisations who provided feedback during the various stages of consultation and in conference presentations.

Introduction and background

The need for international Guidelines

The focus of the International Test Commission (ITC) project is on guidelines for good test use and for encouraging best practice in assessment. The work so far carried out by the ITC to promote good practice in test adaptations (Hambleton, 1994; Van de Vijver, F. & Hambleton, R., 1996) is an important step towards assuring uniformity in the quality of tests adapted for use across different cultures and languages. At its meeting in Athens in 1995, the ITC Council accepted a proposal to broaden this concern to include guidelines on the fair and ethical use of tests, from which standards for training and specifying the competence of test users could be derived.

There are a number of reasons why guidelines on test use are needed at an international level.

- Countries differ greatly in the degree, if any, of statutory control they can exercise over the use of testing and its consequences for those tested. Some national professional societies have statutory registration of psychologists, others do not; some have mechanisms for the control of standards of test use by non-psychologists, others do not. The existence of a set of internationally-accepted guidelines would provide national psychological associations and other relevant professional bodies and organisations with a degree of support in the endeavours of such organisations to develop standards in countries where such standards are currently either lacking in some respect or non-existent.
- Patterns of access, in terms of the rights to purchase or use test materials, vary greatly from country to country. In some countries, access is restricted to psychologists, in others to users registered with formally approved national test distributors, in yet others, test users may be free to obtain materials without restriction from suppliers in their country or directly from suppliers abroad.
- A number of well-known instruments have appeared on the Internet in violation of copyright, without acknowledgement of the test authors or publishers, and without regard to issues of test security.
- Within the occupational testing arena, the greater international mobility of labour has increased the demand for tests to be used on job applicants from a number of different countries - often with the tests being administered in one country on behalf of a potential employer in another country.
- Development work is being carried out in the USA and in the UK on the use of Internet for distance- or remote-assessment in both occupational and educational settings. This raises a whole host of issues relating to standards of administration and control over the testing process, including test security.

Aim and objectives

The long-term aim of this project includes the production of a set of guidelines that relate to the competencies (knowledge, skills, abilities and other personal characteristics) needed by test users. These competencies are specified in terms of assessable performance criteria. These criteria provide the basis for developing specifications of the evidence of competence that would be expected from someone seeking qualification as a test user. Such competencies need to cover such issues as:

- professional and ethical standards in testing,
- rights of the test taker and other parties involved in the testing process,
- choice and evaluation of alternative tests,
- test administration, scoring and interpretation,
- report writing and feedback.

Insofar as they directly relate to test use, the Guidelines also have implications for:
- standards for test construction,
- standards for user-documentation - e.g., technical and user manuals,
- standards for regulating the supply and availability of tests and information about tests.

The present Guidelines represent the work of specialists in psychological and educational testing (i.e. psychologists, psychometricians, test publishers and test developers) drawn from a number of countries. The intention of this document is not to 'invent' new guidelines, but to draw together the common threads that run through existing guidelines, codes of practice, standards and other relevant documents, and to create a coherent structure within which they can be understood and used.

Development of the Guidelines

The Guidelines should be considered as benchmarks against which existing local standards can be compared for coverage and international consistency. By using the Guidelines as benchmarks or the basis from which to develop locally applicable documents (e.g. standards, codes of practice, statements on test taker rights), a high level of consistency across national boundaries will be promoted.

Work on the Guidelines began by drawing together materials concerned with test standards, codes of practice, test use, etc., from a number of countries[1]. While drawing on all of these sources, the present Guidelines have been particularly influenced by:
- The Australian Psychological Society (APS) *Supplement to guidelines on the use of Psychological Tests* (Kendall et al., 1997).
- The British Psychological Society (BPS) *Level A and Level B standards for occupational test use* (Bartram, 1995, 1996).
- The American Educational Research Association (AERA) , American Psychological Association (APA), & National Council on Measurement in Education (NCME) (1985) *Standards for educational and psychological testing.*
- American Association for Counseling and Development (AACD) *Responsibilities of Users of Standardized Tests* (Schafer, W.D, 1992).
- The CPA (Canadian Psychological Association, 1987) *Guidelines for Educational and Psychological Testing*.

The APS document has been particularly valuable as it pulls together much of what is contained in the BPS and American publications as well as drawing on South African National Institute for Psychological Research (NIPR) publications and various publishers' guidance for test users. It also embodies much of what has come out of Joint Committee on Testing Practices (JCTP) Test User Qualifications Working Group's (TUQWG) seminal work on a data-based approach to promoting good test

[1] A list of all the materials that informed this process is available on request from the authors.

use (e.g., Eyde et al, 1988, 1993; Moreland et al., 1995), and the work of the JCTP on the *Code of Fair Testing Practices in Education* (JCTP, 1988; Fremer, Diamond, & Camara, 1989).

The content of the primary sources was analysed and statements categorised under 14 main headings. Where appropriate single statements were written to capture the common meaning of a number of statements from different sources. Statements were also modified in format such that they provided completion of a common stem (e.g.: "Competent test users will endeavour to....", or "Competent test users can....").

This initial structure of 14 main sections and their content was embodied in the first draft Framework Document. This formed the material for an international workshop held in Dublin in July 1997. The purpose of the ITC Workshop was to consider and critically evaluate all aspects of a framework document, with a view to producing a draft set of guidelines that would have international currency and acceptance. During the workshop, the Framework Document was examined in detail, with refinements being proposed in terms of form, structure and content. Following the workshop, the document was extensively revised (Version 2.0) and circulated to all those who attended for comment. A draft consultation document (Version 3.1) was prepared that incorporated all the comments and suggestions submitted regarding Version 2.0.

Copies of the Version 3.1 consultation document and a structured response form were widely circulated to key individuals and organisations for comment. A total of 200 were distributed. A total of 28 detailed responses were received including 'corporate' responses from the APA, the BPS and some other European professional associations. In the summer of 1998 the Guidelines were revised in the light of these comments, and 200 copies (Version 4.1) were sent out for further consultation. A total of 18 formal responses were received to this second round of consultation. In addition, supportive informal comments were provided by many recipients of the consultation documents by email or in meetings.

In producing the current version of the Guidelines (Version 2000), every effort has been made to take account of all these responses. Without exception, the responses were helpful and constructive.[2]

These Guidelines are to be seen as supportive rather than constraining. We need to ensure that the Guidelines embody universal key principles of good test use, without attempting to impose uniformity on legitimate differences in function and practice between countries or between areas of application.

The proposed structure differentiates three main aspects of competence:

1. Professional and ethical standards of good practice that affect the way in which the process of testing is carried out and the way in which test users interact with others involved in the process.
2. The knowledge, understanding and skills relating to the process of testing: what test users need to be able to do.

[2] A detailed report on the results of the first consultation was submitted to the ITC Council meeting in August, 1998. A report on the second consultation together with Version 5.0 of the Guidelines was submitted to the ITC Council when it met in June 1999. Version 2000 contains some minor editorial amendments to Version 5.0.

3. The knowledge and understanding that are necessary to inform and underpin the process of testing.

These three components differ, yet are inextricably inter-twined in practice.
The Guidelines start from a *key purpose.* This can be characterised as the 'mission statement' for test users. It provides the focus from which the guidelines are developed. Each guideline defines an aspect of test user competence that contributes to the key purpose.

Together with the key purpose, the *scope statement* describes to whom the Guidelines apply, the forms of assessment to which they relate, and the assessment contexts.

This document contains:

1. Key purpose and scope statements.
2. Specifications of test user competencies in relation to ethical test use.
3. Specifications of test user competencies in relation to good practice in the use of tests.

The Guidelines

Key purpose

A competent test user will use tests appropriately, professionally, and in an ethical manner, paying due regard to the needs and rights of those involved in the testing process, the reasons for testing, and the broader context in which the testing takes place.

This outcome will be achieved by ensuring that the test user has the necessary competencies to carry out the testing process, and the knowledge and understanding of tests and test use that inform and underpin this process.

Scope of the Guidelines

Any attempt to provide a precise definition of a 'test' or of 'testing' as a process, is likely to fail as it will tend to exclude some procedures that should be included and include others that should be excluded. For the purpose of these Guidelines, the terms 'test' and 'testing' should be interpreted broadly. Whether an assessment procedure is labelled a 'test' or not is immaterial. These Guidelines will be relevant for many assessment procedures that are not called 'tests' or that seek to avoid the designation 'test'. Rather than provide a single definition, the following statements attempt to map out the domain covered by the Guidelines.

- Testing includes a wide range of procedures for use in psychological, occupational and educational assessment.
- Testing may include procedures for the measurement of both normal and abnormal or dysfunctional behaviours.
- Testing procedures are normally designed to be administered under carefully controlled or standardised conditions that embody systematic scoring protocols.
- These procedures provide measures of performance and involve the drawing of inferences from samples of behaviour.
- They also include procedures that may result in the qualitative classification or ordering of people (e.g., in terms of type).

Any procedure used for 'testing', in the above sense, should be regarded as a 'test', regardless of its mode of administration; regardless of whether it was developed by a professional test developer; and regardless of whether it involves sets of questions, or requires the performance of tasks or operations (e.g., work samples, psycho-motor tracking tests).

Tests should be supported by evidence of reliability and validity for their intended purpose. Evidence should be provided to support the inferences that may be drawn from the scores on the test. This evidence should be accessible to the test user and available for independent scrutiny and evaluation. Where important evidence is contained in technical reports that are difficult to access, fully referenced synopses should be provided by the test distributor.

The test use Guidelines presented here should be considered as applying to all such procedures, whether or not they are labelled as 'psychological tests' or 'educational tests' and whether or not they are adequately supported by accessible technical evidence.

Many of these Guidelines will apply also to other assessment procedures that lie outside the domain of 'tests'. They may be relevant for any assessment procedure that is used in situations where the assessment of people has a serious and meaningful intent and which, if misused, may result in personal loss or psychological distress (for example, job selection interviews, job performance appraisals, diagnostic assessment of learning support needs).

The Guidelines do not apply to the use of materials that may have a superficial resemblance to tests, but which all participants recognise are intended to be used only for purposes of amusement or entertainment (e.g., life-style inventories in magazines or newspapers).

Who the Guidelines are for

The Guidelines apply to the use of tests in professional practice. As such they are directed primarily towards:

- The purchasers and holders of test materials;
- Those responsible for selecting tests and determining the use to which tests will be put;
- Those who administer, score or interpret tests;
- Those who provide advice to others on the basis of test results (e.g., recruitment consultants, educational and career counsellors, trainers, succession planners);
- Those concerned with the process of reporting test results and providing feedback to people who have been tested.

The Guidelines will be of relevance to others involved in the use of tests as defined above. These include:

- the developers of tests,
- the suppliers of tests,
- those involved in the training of test users,
- those who take tests and their relevant others (e.g., parents, spouses, partners),
- professional bodies and other associations with an interest in the use of psychological and educational testing, and
- policy makers and legislators.

While aimed primarily at professional practice, most aspects of the good practice embodied in the Guidelines will also be of relevance to those who use tests solely for research purposes.

The Guidelines are not intended to cover every type of assessment technique (e.g., unstructured or semi-structured interviews, assessed group activities) or every situation in which assessment occurs (e.g., employment assessment centres). Yet

many of the Guidelines are likely to be applicable in assessment situations and for purposes more general than those concerned primarily with psychological and educational testing (for example, the use of assessment centres for employmee placement or selection, semi-structured and structured interviews, or assessment for selection, career guidance and counselling).

Contextual factors

The Guidelines are applicable internationally. They may be used to develop specific local standards through a process of contextualisation. It is recognised that there are many factors which affect how standards may be managed and realised in practice. These contextual factors have to be considered at the local level when interpreting the Guidelines and defining what they would mean in practice within any particular setting.

The factors that need to be considered in turning Guidelines into specific standards include:

- social, political, institutional, linguistic, and cultural differences between assessment settings;
- the laws of the country in which testing is taking place;
- existing national guidelines and performance standards set by professional psychological societies and associations;
- differences relating to individual versus group assessment;
- differences related to the test setting (educational, clinical, work-related and other assessment);
- who the primary recipients of the test results are (e.g., the test-takers, their parents or guardian, the test-developer, an employer or other third party);
- differences relating to the use of test results (e.g., for decision-making, as in selection screening, or for providing information to support guidance or counselling); and
- variations in the degree to which the situation provides opportunity for the accuracy of interpretations to be checked in the light of subsequent information and amended if needed.

Knowledge, Understanding, and Skill

Knowledge, understanding and skill underpin all the test user competencies. The nature of their content and level of detail are likely to vary between countries, areas of application and as a function of the level of competence required to use a test.

The Guidelines do not contain detailed descriptions of these. However, when applying the Guidelines for use in specific situations the relevant knowledge, skills, abilities and other personal characteristics will need to be specified. This specification is part of the process of contextualisation, through which generic guidelines are developed into specific standards. The main areas descriptions of knowledge, understanding and skills need to cover include the following.

Relevant declarative knowledge.
This includes:

- knowledge of basic psychometric principles and procedures, and the technical requirements of tests (e.g., reliability, validity, standardisation);
- knowledge of tests and measurement sufficient to enable the proper understanding of test results;
- knowledge and understanding of relevant theories and models of ability, of personality or other psychological constructs, or of psychopathology, as necessary to properly inform the choice of tests and the interpretation of test results; and
- knowledge of the tests and the test suppliers relevant to one's area of practice.

Instrumental knowledge and skills
These include:

- knowledge and skills relating to specific assessment procedures or instruments, including the use of computer-based assessment procedures;
- specialised knowledge of and practitioner skills associated with using those tests that are within one's repertoire of assessment tools; and
- knowledge and understanding of the construct or constructs underlying test scores, where this is important if valid inferences are to be drawn from the test results.

The Guidelines cover:

General personal task-related skills

This includes:

- the performance of relevant activities such as test administration, reporting, and the provision of feedback to test takers and other clients;
- oral and written communication skills sufficient for the proper preparation of test takers, test administration, the reporting of test results, and for interaction with relevant others (e.g., parents, or organisational policy makers); and
- interpersonal skills sufficient for the proper preparation of test takers, the administration of tests, and the provision of feedback of test results.

Contextual knowledge and skills

This includes:

- knowing when and when not to use tests;
- knowing how to integrate testing with other less formal components of the assessment situation (e.g., biographical data, unstructured interview and references etc.); and
- knowledge of current professional, legal, and ethical issues relating to the use of tests, and of their practical implications for test use.

Task management skills

This includes:

- knowledge of codes of conduct and good practice relating to the use of tests, test

data, the provision of feedback, the production and storage of reports, the storage of and responsibility for test materials and test data; and
- knowledge of the social, cultural, and political context in which the test is being used, and the ways in which such factors might affect the results, their interpretation and the use to which they are put.

Contingency management skills

This includes:
- knowing how to deal with problems, difficulties, and breakdowns in routine;
- knowing how to deal with a test taker's questions during test administration etc.; and
- knowing how to deal with situations in which there is the potential for test misuse or for misunderstanding the interpretation of test scores.

1 Take responsibility for ethical test use

Competent test users should:

1.1 Act in a professional and ethical manner

1.1.1 Promote and maintain professional and ethical standards.
1.1.2 Have a working understanding of current professional and ethical issues and debates relating to the use of tests in their field of application.
1.1.3 Implement an explicit policy on testing and test use.[3]
1.1.4 Ensure that people who work for or with them adhere to appropriate professional and ethical standards of behaviour.
1.1.5 Conduct communications with due concern for the sensitivities of the test taker and other relevant parties.
1.1.6 Represent tests and testing in a positive and balanced manner in communications with and through the media.
1.1.7 Avoid situations in which they may have or be seen to have a vested interest in the outcome of the assessment, or where the assessment might damage the relationship with their client.

1.2 Ensure they have the competence to use tests

1.2.1 Work within the limits of scientific principle and substantiated experience.
1.2.2 Set and maintain high personal standards of competence.
1.2.3 Know the limits of their own competence and operate within those limits.
1.2.4 Keep up with relevant changes and advances relating to the tests they use, and to test development, including changes in legislation and policy, which may impact on tests and test use.

1.3 Take responsibility for their use of tests

1.3.1 Only offer testing services and only use tests for which they are qualified.
1.3.2 Accept responsibility for the choice of tests used, and for the recommendations made.
1.3.3 Provide clear and adequate information to participants in the testing process about the ethical principles and legal regulations governing psychological testing.
1.3.4 Ensure that the nature of the contract between test-taker and tester is clear and understood.[4]
1.3.5 Be alert to any unintended consequences of test use.
1.3.6 Endeavour to avoid doing harm or causing distress to those involved in the

[3] An example policy outline is attached as Appendix A.

[4] An example 'contract' between test user and test taker is attached as Appendix B.

testing process.

1.4 Ensure that test materials are kept securely

1.4.1 Ensure secure storage of and control access to test materials
1.4.2 Respect copyright law and agreements that exist with respect to a test including any prohibitions on the copying or transmission of materials in electronic or other forms to other people, whether qualified or otherwise.
1.4.3 Protect the integrity of the test by not coaching individuals on actual test materials or other practice materials that might unfairly influence their test performance.
1.4.4 Ensure that test techniques are not described publicly in such a way that their usefulness is impaired

1.5 Ensure that test results are treated confidentially.

1.5.1 Specify who will have access to results and define levels of confidentiality.
1.5.2 Explain levels of confidentiality to individuals before tests are administered .
1.5.3 Limit access to results to those with a right to know.
1.5.4 Obtain the relevant consents before releasing results to others.
1.5.5 Protect data kept on file so that only those who have a right of access can obtain them.
1.5.6 Establish clear guidelines as to how long test data are to be kept on file.
1.5.7 Remove names and other personal identifiers from databases of results that are archived, for research use, development of norms or other statistical purposes.

2 Follow good practice in the use of tests

2.1 Evaluate the potential utility of testing in an assessment situation

Competent test users will:

2.1.1 Produce a reasoned justification for the use of tests.

2.1.2 Ensure there has been a thorough analysis of the client's needs, reasons for referral, or of the diagnostic category, condition, or job for which assessment is being used.

2.1.3 Establish that the knowledge, skills, abilities, aptitudes or other characteristics, which the tests are intended to measure, are correlates of relevant behaviours in the context about which inferences are to be drawn.

2.1.4 Seek other relevant collateral sources of information.

2.1.5 Assess the advantages and disadvantages of using tests compared with other sources of information.

2.1.6 Ensure that full use is made of all available collateral sources of information.

2.2 Choose technically sound tests appropriate for the situation

Competent test users will:

2.2.1 Examine current information covering the range of potentially relevant tests (e.g., from specimen sets, independent reviews, expert advice), before selecting a test to use.

2.2.2 Determine that the test's technical and user documentation provides sufficient information to enable evaluation of the following:

- a) scope or coverage and representativeness of test content, appropriateness of norm groups, difficulty level of content etc.;
- b) accuracy of measurement and reliability demonstrated with respect to relevant populations;
- c) validity (demonstrated with respect to relevant populations) and relevance for the required use;
- d) freedom from systematic bias in relation to the intended test taker groups;
- e) acceptability to those who will be involved in their use, including perceived fairness and relevance;
- f) practicality, including time required, costs, and resource needs.

2.2.3 Avoid the use of tests that have inadequate or unclear supporting technical documentation;

2.2.4 Use tests only for those purposes where relevant and appropriate validity evidence is available.

2.2.5 Avoid judging a test solely on the basis of face value, test-user testimonials, or advice from those with a vested commercial interest.

2.2.6 Respond to requests from relevant interested parties (e.g. test takers, parents, managers) by providing sufficient information to allow them to understand why the test was chosen.

2.3 Give due consideration to issues of fairness in testing

When tests are to be used with individuals from different groups (e.g., groups differing in terms of gender, cultural background, education, ethnic origin, or age), competent test users will make all reasonable efforts to ensure that:

2.3.1 The tests are unbiased and appropriate for the various groups that will be tested.
2.3.2 The constructs being assessed are meaningful in each of the groups represented.
2.3.3 Evidence is available on possible group differences in performance on the test.
2.3.4 Evidence relating to differential item functioning (DIF) is available, where relevant.
2.3.5 There is validity evidence to support the intended use of the test in the various groups.
2.3.6 Effects of group differences not relevant to the main purpose (e.g., differences in motivation to answer, or reading ability) are minimised.
2.3.7 In all cases, Guidelines relating to the fair use of tests are interpreted in the context of local policy and legislation.[5]

When testing in more than one language (within or across countries[6]), competent test users will make all reasonable efforts to ensure that:

2.3.8 Each language or dialect version has been developed using a rigorous methodology meeting the requirements of best practice.
2.3.9 The developers have been sensitive to issues of content, culture and language.
2.3.10 The test administrators can communicate clearly in the language in which the test is to be administered.
2.3.11 The test taker's level of proficiency in the language in which the test will be administered is determined systematically and the appropriate language version is administered or bilingual assessment is performed, if appropriate.

When tests are to be used with people with disabilities, competent test users will make all reasonable efforts to ensure that:

2.3.12 Advice is sought from relevant experts on the potential effects of the various disabilities on test performance.
2.3.13 Potential test takers are consulted and their needs and wishes are given proper consideration.
2.3.14 Adequate arrangements are made when test takers include people with hearing, visual or motor impairments, or other disabilities (e.g., learning impairments, dyslexia .).
2.3.15 Use of alternative assessment procedures, rather than modifications to tests, is considered (e.g., other more suitable tests, or alternative structured forms of assessment).
2.3.16 Relevant professional advice is sought if the degree of modification required

[5] The Guidelines in this section focus on what is 'best practice'. However, in many countries, issues relating to the fair use of tests must also take account of national laws (e.g., the Americans with Disabilities Act, 1990, in the USA, or the Race Relations Act, 1976, in the UK).
[6] These Guidelines relate not only to different national languages and dialects, but also to special forms of communication, such as sign language, used to overcome the effects of forms of disability.

for use by those with disabilities is beyond the experience of the test user.

2.3.17 Modifications, when necessary, are tailored to the nature of the disability and are designed to minimize impact on score validity.

2.3.18 Information regarding the nature of any modifications made to a test or testing procedure is provided to those who interpret or act upon the test scores whenever the withholding of such information might otherwise result in biased interpretation or an unfair decision.

2.4 Make necessary preparations for the testing session

The competent test user will make all reasonable efforts to:

2.4.1 Provide relevant parties in a timely manner with clear information concerning the purpose of testing, ways in which they might best prepare for the test session, and the procedures to be followedr.

2.4.2 Advise test takers of the linguistic or dialectic groups for which the test is considered appropriate.

2.4.3 Send test takers approved practice, sample, or preparation materials where these are available and where this is consistent with recommended practice for the tests concerned.

2.4.4 Explain clearly to test takers their rights and responsibilities[7].

2.4.5 Gain the explicit consent of test takers or their legal guardians or representatives before any testing is done.

2.4.6 Explain, when testing is optional, the consequences of taking or not taking the test to relevant parties so that they can make an informed choice.

2.4.7 Make the necessary practical arrangements by ensuring that:

a) preparations conform to those stipulated in the publisher's manual;

b) locations and facilities for testing have been arranged well in advance, and the physical environment is accessible, safe, quiet, free from distractions and appropriate for the purpose;

c) sufficient materials are available and have been checked to ensure there are no marks left by previous users on question booklets or answer sheets;

d) staff who will be involved in the administration are competent;

e) appropriate arrangements have been made for the testing of people with disabilities[8].

2.4.8 Anticipate likely problems and counteract them through thorough preparation of materials and instructions.

2.5 Administer the tests properly

The competent test user will:

2.5.1 Establish rapport by welcoming test-takers and briefing them in a positive fashion.

2.5.2 Act to reduce test-taker anxiety and avoid creating or reinforcing unnecessary anxiety.

2.5.3 Ensure potential sources of distraction (e.g., wristwatch alarms, mobile

[7] See Appendix B.
[8] See Appendix C.

phones, pagers) are removed.
2.5.4 Ensure test-takers have the materials they require for taking the test before it begins.
2.5.5 Administer tests under appropriate supervised conditions.
2.5.6 Wherever possible, administer test instructions in the primary language of the test takers, even where the test content is designed to provide evidence of knowledge or skills in a non-primary language.
2.5.7 Adhere strictly to the directions and instructions as specified in the test manual while making reasonable accommodations for persons with disabilities.
2.5.8 Read instructions clearly and calmly.
2.5.9 Provide adequate time for examples to be completed.
2.5.10 Observe and record deviations from test procedures.
2.5.11 Monitor and record response times accurately where appropriate.
2.5.12 Ensure all materials are accounted for at the end of each testing session.
2.5.13 Administrer tests by modes that permit adequate and appropriate levels of supervision and authentication of the identity of the test takers.
2.5.14 Ensure those assisting the administration have had proper training.
2.5.15 Ensure test takers are not left unattended or subjected to distracting activities during a supervised test session.
2.5.16 Provide appropriate assistance to test takers who show signs of undue distress or anxiety.

2.6 Score and analyse test results accurately

Competent test users will:
2.6.1 Follow carefully the standardised procedures for scoring.
2.6.2 Carry out appropriate raw score conversions to other relevant types of scale.
2.6.3 Choose scale types relevant to the intended use of the test scores.
2.6.4 Check score scale-conversions and other clerical procedures for accuracy.
2.6.5 Ensure that invalid conclusions are not drawn from comparisons of scores with norms that are not relevant to the people being tested or are outdated.
2.6.6 Compute, where appropriate, composite scores using standard formulae and equations.
2.6.7 Employ procedures to screen test results to recognise improbable or unreasonable scores.
2.6.8 Clearly and accurately label scales in reports, and provide clear identification of norms, scales types, and equations used.

2.7 Interpret results appropriately

Competent test users will:
2.7.1 Have a good professional understanding of the test's theoretical or conceptual basis, technical documentation and guidance on the use and interpretation of the scale scores.
2.7.2 Have a good understanding of the scales used, the characteristics of the norm or comparison groups, and the limitations of the scores.
2.7.3 Take steps to minimise the effects on test interpretation of any biases the test interpreter may have towards members of the test taker's cultural group.
2.7.4 Use appropriate norm or comparison groups where available.
2.7.5 Interpret results in the light of available information about the person being

tested (including age, gender, schooling, culture and other factors) with due regard for the technical limitations of the test, the assessment context, and the needs of those with a legitimate interest in the outcome of the process.

2.7.6 Avoid over-generalising the results of one test to traits or human characteristics which are not measured by the test.

2.7.7 Consider each scale's reliability, error of measurement and other qualities which may have artificially lowered or raised results when interpreting scores.

2.7.8 Give due consideration to the available evidence of validity, with respect to the construct being measured for members of the test takers' relevant demographic groups (e.g., cultural, age, social class, and gender groups).

2.7.9 Use passing scores (cut-scores) in test interpretation only when evidence of the validity for the pass scores is available and supports its use.

2.7.10 Be aware of negative social stereotyping that may pertain to members of the test taker's group (e.g., cultural group, age, social class, and gender) and avoid interpreting tests in a manner that perpetuates such stereotyping.

2.7.11 Take into account any individual or group variations from standard procedures in test administration.

2.7.12 Take into account any evidence of prior experience with the test where there are data available relating to the effect of such experience on test performance.

2.8 Communicate the results clearly and accurately to relevant others

Competent test users will:

2.8.1 Identify appropriate parties who may legitimately receive test results.

2.8.2 With the informed consent of the test takers, or their legal representatives, produce written or oral reports for relevant interested parties.

2.8.3 Ensure that the technical and linguistic levels of any reports are appropriate for the level of understanding of the recipients.

2.8.4 Make clear that the test data represent just one source of information and should always be considered in conjunction with other information.

2.8.5 Explain how the importance of the test results should be weighted in relation to other information about the people being assessed.

2.8.6 Use a form and structure for a report that is appropriate to the context of the assessment.

2.8.7 When appropriate, provide decision-makers with information on how results may be used to inform their decisions.

2.8.8 Explain and support the use of test results used to classify people into categories (e.g., for diagnostic purposes or for job selection).

2.8.9 Include within written reports a clear summary, and when relevant, specific recommendations.

2.8.10 Present oral feedback to test takers in a constructive and supportive manner.

2.9 Review the appropriateness of the test and its use

Competent test users will:

2.9.1 Monitor and periodically review changes over time in the populations of individuals being tested and any criterion measures being used.

2.9.2 Monitor tests for evidence of adverse impact.

2.9.3 Be aware of the need to re-evaluate the use of a test if changes are made to its

form, content, or mode of administration.

2.9.4 Be aware of the need to re-evaluate the evidence of validity if the purpose for which a test is being used is changed.

2.9.5 Where possible, seek to validate tests for the use to which they are being put, or participate in formal validation studies.

2.9.6 Where possible, assist in updating information regarding the norms, reliability and validity of the test by providing relevant test data to the test developers, publishers or researchers.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for Educational and Psychological Testing.* Washington DC: American Psychological Association.

Bartram, D. (1995). The Development of Standards for the Use of Psychological Tests in Occupational Settings: The Competence Approach. *The Psychologist*, May, 219-223.

Bartram, D. (1996). Test Qualifications and Test Use in the UK: The Competence Approach. *European Journal of Psychological Assessment,* **12**, 62-71.

Canadian Psychological Association. (1987). *Guidelines for Educational and Psychological Testing.* Ottawa: Canadian Psychological Association.

Eyde, L. D., Moreland, K. L. & Robertson, G. J. (1988). *Test User Qualifications: A Data-based Approach to Promoting Good Test Use.* Report for the Test User Qualifications Working Group. Washington DC: American Psychological Association.

Eyde, L. D., Robertson, G. J., Krug, S. E. et al (1993*). Responsible Test Use: Case Studies For Assessing Human Behaviour.* Washington DC: American Psychological Association.

Fremer, J., Diamond, E. E. & Camara, W. J. (1989). Developing a Code of Fair Testing Practices in Education. *American Psychologist,* **44**, 1062-1067.

Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, **10***, 229-244.

Joint Committee on Testing Practices. (1988). *Code of Fair Testing Practices in Education*. Washington DC: Joint Committee on Testing Practices.

Kendall, I., Jenkinson, J., De Lemos, M. & Clancy, D. (1997). *Supplement to Guidelines for the use of Psychological Tests.* Australian Psychological Society.

Moreland, K. L., Eyde, L. D., Robertson, G. J., Primoff, E. S. & Most, R. B. (1995). Assessment of Test User Qualifications: A Research-Based Measurement Procedure. *American Psychologist,* **50**, 14-23.

Schafer, W. D. (1992*). Responsibilities of Users of Standardized Tests: RUST Statement Revised.* Alexandria, VA: American Association for Counseling and Development.

Van de Vijver, F. & Hambleton, R. (1996). Translating tests: some practical guidelines. *European Psychologist , **1,*** 89-99.

Appendix A: Guidelines for an outline policy on testing.

The following guidelines relate to the need for organizations to consider their policy on testing in a systematic manner and to ensure that everyone involved is clear as to what the policy is. The need for an explicit policy on testing is not confined to large organisations. Small and medium-sized enterprises that use testing, as well as large ones, should pay regard to testing policy in the same way as they do to health and safety, equal opportunities, disability and other areas relating to good practice in the management, treatment and care of personnel.

While the following considerations or requirements may need to be adapted for use by individual test users operating as sole professional practitioners, it remains important that they have a clear understanding of their own policy and can communicate it to others.

A policy on testing is produced in order to:
- ensure personal and organisational aims are met;
- ensure that potential misuse is avoided;
- demonstrate commitment to good practice;
- ensure test use is appropriate for its purpose;
- ensure tests do not discriminate unfairly;
- ensure evaluations are based on comprehensive, relevant information;
- ensure tests are only used by qualified staff.

A policy on testing will need to cover most if not all the following issues:
- proper test use;
- security of materials and scores;
- who can administer tests, score and interpret tests;
- qualification requirements for those who will use the tests;
- test user training;
- test taker preparation;
- access to materials and security;
- access to test results and test score confidentiality issues;
- feedback of results to test takers;
- responsibility to test takers before, during and after test session;
- responsibilities & accountability of each individual user.

Any policy needs to be regularly reviewed and updated as advances in testing, or changes in practice occur.

Relevant parties need to have access to and be informed about the policy on testing.

Responsibility for any organisation’s testing policy should reside with a qualified test user who has the authority to ensure implementation of and adherence to the policy.

Appendix B: Guidelines for developing contracts between parties involved in the testing process.

Contracts between the test user and test takers should be consistent with good practice, legislation and the test user's policy on testing. The following is provided as an example of the sort of matters such a contract might cover. The details will vary as a function of the assessment context (e.g., occupational, educational, clinical, forensic) and local or national regulations and laws.

Contracts between test user, test takers and other parties are often implicit and unspoken (at least in part). Making clear the expectations, roles and responsibilities of all parties can help to avoid misunderstanding, harm, and litigation.

For their part, the test user will endeavour to:

b.1 inform test takers of their rights regarding how their test scores will be used and their rights of access to them[9];

b.2 give adequate prior warning of any financial charges that may be entailed by the testing process, who will be responsible for their payment, and when payment will be due;

b.3 treat test takers with courtesy, respect and impartiality regardless of race, gender, age, disability, etc.;

b.4 use tests of proven quality, appropriate for the test takers, and appropriate for the assessment purpose;

b.5 inform test takers prior to testing about the purpose of the assessment, the nature of the test, to whom test results will be reported and the planned use of the results;

b.6 give advance notice of when the test will be administered, and when results will be available, and whether or not test takers or others may obtain copies of the test, their completed answer sheets, or their scores[10];

b.7 have a trained person administer the test and have the results interpreted by a qualified person;

b.8 ensure test takers know if a test is optional and, when it is, the consequences of taking or not taking the test;

b.9 ensure test takers understand the conditions, if any, under which they may re-take tests, have tests re-scored, or have their scores cancelled;

b.10 ensure test takers know that they will have their results explained to them as soon as possible after taking the test in easily understood terms;

b.11 ensure test takers understand that their results are confidential to the extent allowed by law and best practice;

b.12 inform test takers who will have access to their results, and the conditions which scores will be released;

b.13 ensure that test takers are aware of the procedures for making complaints or notifying problems;

---

[9] Legislation varies between countries on this issue. For example, the current UK Data Protection Act provides rights of access to data stored on computer different from those for data written on paper.

[10] While tests and answer sheets are not normally passed on to others, there is some variation between countries in practice relating to what test takers or others are permitted to have. However, there is much greater variation in the expectations of test takers concerning what information they will be given. It is important that contracts make clear what they will *not* be given as well as what they will.

The test user will inform test-takers that they are expected to:

b.14 treat others with courtesy and respect during the testing process;

b.15 ask questions prior to testing if uncertain about why the test is to be administered, how it will be administered, what they will be required to do and what will be done with the results;

b.16 inform an appropriate person about any condition that they believe might invalidate the test results or which they would wish to have taken into consideration;

b.17 follow the instructions of the test administrator;

b.18 be aware of the consequences of not taking a test if they choose not to take it, and be prepared to accept those consequences;

b.19 ensure that, if required to pay for any the testing service(s), payment is made by the agreed date.

Appendix C: Points to consider when making arrangements for testing people with disabilities or impairments

Considerable care and expertise is needed when the mode of administration of a test has to be changed to accommodate the needs of people with disabilities. As always, local and national law and practice[11] need to be considered, and the individual's rights to privacy must be respected. In seeking information regarding types and levels of disability, inquiries should only seek information relating to each person's ability to undertake the activities required to complete the test. Particular care needs to be exercised in relation to employment testing[12].

There is no simple rule of thumb that can be used to ensure that a test is administered fairly for people with all types of disability. It is a matter of professional judgement as to whether it is better to use some alternative form of assessment, or to modify the test or its mode of administration. In practice, it is rarely possible to norm modified tests on sufficient samples of people with equivalent disability in order to ensure comparability of the test with the standardised version. However, where data exist on, for example, the effects of changing time limits, use of Braille or audiotape spoken versions of tests, such data should guide the user in making the necessary accommodations. While full standardization of a modified version may not be possible, pilot testing on small samples of individuals should be carried out whenever practical.

Given the dearth of information about the performance of people with disabilities on tests (whether modified or not), it is often more appropriate for test result to be used in a more qualitative manner. They can be used to give an indication of the characteristic being assessed (ability, motivation, personality, etc.), which can be supplemented and supported by information gathered using other methods.

For individual assessment, the assessor can usually tailor the assessment procedures to the capabilities of the person who is being assessed. However, particular issues arise in group testing (e.g., for selection into employment). Here there may be practical difficulties involved in varying the mode of administration for particular individuals within a group administration setting. Furthermore, all parties may see differences in treatment as being unfair. For example if more time is given for test completion, those with the disability may be conscious that they are being treated 'differently', and those without the disability may feel that the extra time provides an unfair advantage.

Advice on special needs can usually be obtained from relevant disability organisations as well as the individual test takers. It is generally helpful (where the law permits) to ask the individual directly in a non-threatening and supportive way if there are any considerations that need to be taken into account[13]. In many cases such consultation

---

[11] In the United States, for example, attention must be paid to the provisions of the Americans with Disabilities Act (1990). In the UK, the Disability Discrimination Act (1995), Employment Code of Practice states that "employers are required to revise tests – or the way the results of such tests are assessed – to take account of specific disabled candidates."

[12] For detailed guidance on this in the United States, see Eyde, Nestor, Heaton and Nelson (1994).

[13] In the UK, the Disability Discrimination Act (1995) also places some obligation on the individual to raise awareness of their needs.

will enable suitable modifications to be made to the test taking environment without requiring changes to the test itself.

The following outline protocol provides a general guide to the process of deciding whether to modify testing and how to carry out the modification. Essentially, disability may contribute no variance to test scores, contribute construct relevant variance or construct irrelevant variance. In the first case, no modifications are necessary. In the final case, modifications should be aimed at removing the irrelevant source of variance (by suitable modification of the test conditions or substitution of a more suitable test). For the second case (construct relevant variance), however, modification to the test will affect the relevance of the test scores.

c1. Is the disability likely to have an effect upon test performance? Many people have disabilities that would not affect test performance. In such cases, it would be inappropriate to make accommodations for them.
c2. If the disability is likely to affect test performance, then is the effect on performance incidental to the construct being measured? For example, a person with an arthritic hand may have trouble with a speeded test which involves writing. If the ability to perform manual tasks rapidly is part of the construct being measured, then the test should not be changed. However, if the purpose is to assess visual checking speed, then an alternative mode of response would be appropriate.
c3. When the particular disability is incidental to the construct being measured but is likely to affect the individual's performance on the test, then modification of the procedure may be considered.
c4. Users should always consult the test manual and the publisher for guidance on modification and for information regarded alternative formats and procedures.
c5. Users should also consult relevant disability organisations for advice and guidance on the possible implications of a specific disability, relevant literature or documentation, and the sort of adaptations or accommodations that may prove helpful.
c6. Any modifications made to the test or test administration procedures should be carefully documented along with the rationale behind the modification.

