EFPA

WORKING GROUP ON
TEST USER STANDARDS
Convenor : DAVE BARTRAM

REPORT to the
GENERAL ASSEMBLY 2005 in GRANADA

EWG1

Report of the joint EFPA-EAWOP European Working Group (EWG) on the development of a European standard for test user qualification in the field of work and organisational assessment

Prepared for the EFPA and EAWOP General Assemblies by

The EWG co-convenors:

Dave Bartram (EFPA) & Henry Honkanen (EAWOP)

May 2005

1. EXECUTIVE SUMMARY

The EWG was set up by EFPA and EAWOP in 2003 to develop a set of common European standards for users of psychological tests in the field of work and organizational assessment.

The EWG has built on the work of the International Test Commission (ITC) to produce a set of Standards, which define performance requirements and related learning specifications covering the areas of good practice set out in the ITC's International Guidelines on Test Use. The EWG has also produced illustrative assessment guidance to support the standards. This work has built on the experience of various different projects on test user certification in the UK, the Netherlands, Sweden, Norway and Finland.

The standards and related guidance have been subject to detailed review by members of the EWG and have been the subject of an in-depth European consultation process. Preliminary results from that consultation have been strongly supportive of the project.

The recommendations of the EWG to the General Assemblies of EFPA and EAWOP are that:

1. The standards (see Annex) are recognised by the General Assemblies as an acceptable basis for the further development of European Standards for Test Use in the field of Work and Organizational assessment.
2. Through its Standing Committee on Tests and Testing, EFPA establishes procedures to support and maintain these standards.
3. That EFPA and EAWOP, through their Executive Committees, establish a small joint board responsible for reviewing progress, and dealing with policy issues.
4. EAWOP takes the lead role in implementing a European process for accrediting local national qualifications in test use based on the Standards and for establishing a European Register of Test Users.
5. A small joint EFPA-EAWOP implementation group is established to explore issues relating to the above.
    1. Complete the standards development and definition work in the light of the feedback from the consultation.
    2. Develop a set of rules and regulations regarding European accreditation and European certification, using the European Diploma in Psychology regulations as a model.
    3. Collaborate with current certification providers and pilot projects launched at the national levels.
    4. Prepare a proposal for how the Standards and the accreditation is administrated and managed at the European level, including a definition of the responsible partners and bodies.
    5. Produce detailed recommendations on implementation for the EFPA and EAWOP Executive Committees in 2007 with final proposals for the standards and their implementation.

2 INTRODUCTION

In 2003 the EFPA Executive approved the foundation of a European Working Group (EWG) to explore, jointly with EAWOP, the development of a European standard for test user qualification in the field of work and organizational psychology.

The genesis of this EWG goes back to the 2001 London meeting of the EFPA Standing Committee on Tests and Testing, where it was agreed that we should set up a project to explore the harmonisation or equating of psychological test user qualifications across Europe. We agreed to work most closely with countries that were actively developing this: At that time, this was the UK, Sweden and Netherlands. Subsequently, work has also moved ahead in Norway, in conjunction with the DNV and in Finland. For practical reasons, we agreed to focus our attention on the area of psychological testing in the field of work, rather than health or education.

Since that 2001 meeting a number of significant developments occurred:

1. The STP has been progressing qualifications in Sweden based on the British Psychological Society's (BPS) standards. A number of qualifications have now been awarded. Unlike the BPS Level A and Level B Certification model, in Sweden there is a single qualification (comparable to the British Level A plus the Level B Intermediate).
2. In Norway, DNV and the Norwegian Psychological Association, together with other stakeholders, have developed a structure and process to manage the development and delivery of test user qualifications. This is now progressing well and it is expected to be ready to deliver qualifications later this year.
3. In Finland, the Finnish Psychological Association has set up user qualification procedures.
4. CITO-4TP in consultation with the Dutch Psychological Association (NIP) is progressing a qualification process in Netherlands
5. The DIN 33430 standard in Germany provides an opportunity for links with user qualification procedures.
6. The BPS has been carrying out an overall review of its Level A and Level B test user qualification framework and standards, which are now over 10 years old. To date over 20,000 people in the UK have received BPS certificates of competence in test use since the programme started in 1991.

The role of the EWG

In the light of these developments, at the meeting of the EFPA Standing Committee on Tests and Testing in Vienna in July 2003, it was agreed that the project should do more than simply monitor and observe what was happening in Europe. We agreed to set up a working group. Because of the focus on testing in work and organizational settings, we formally invited EAWOP to join with us as equal partners in this work.

This EWG was constituted to:

1. Be accountable to the EFPA General Assembly through the Standing Committee on Tests and Testing and to the EAWOP General Assembly
2. Collect information and review the work being done by any country that can feed into the definition of standards and evidence requirements for psychological test user qualifications.
3. Carry out a consultation process on this with countries wishing to develop their own qualification
4. Provide a draft set of standards for consideration at the 2004 meeting of the Committee, with a view to submitting these for formal approval at the respective EFPA and EAWOP General Assemblies in 2005.

5. Depending upon the outcome of the work, the EWG will make proposals for future actions after 2005.

The EWG co-convenors are Dave Bartram (UK, for EFPA) and Henry Honkanen (Finland, for EAWOP) and it has active representation from the following countries: Belgium, Denmark, Finland, Germany, Netherlands, Norway, Spain, Sweden, UK.

History of the development of the standards

The EWG began its work with a meeting in London in December 2003. Subsequently it met in Palma in May 2004 and again in Prague in November 2004. Following the Prague meeting there was broad agreement on the form and content of the Standards, and work continued on drafting and restructuring the documentation in time for a final meeting of the group in Brussels in March 2005. Following that meeting, materials were prepared for European-wide consultation. Given the timescale for the consultation, an oral report on the final outcome will be provided to the General Assemblies of EAWOP and EFPA if required. The preliminary results from this consultation have been very supportive and are summarised in the Appendices to this report.

It has worked from two key sets of documentation. The ITC Guidelines on Test Use and the BPS Level A and Level B Test User qualification standards. The latter have been undergoing a process of revision and updating in parallel with the EWG project, and it has been beneficial to both activities to have a very close degree of collaboration between the EWG and the BPS working party. In addition, all those actively developing test user qualifications in Europe for work and organizational test use have been actively involved in the drafting and development of the standards – both through activities during the EWG meetings and in between meetings.

Why Test User Standards are needed

Within the work and organizational field, there is increasing interest in Europe in the development of test user standards and test standards. This interest arises not only from the psychology profession but also from allied professions, like human resources management.

At present there is little if any uniformity across Europe in terms of test user qualification standards or processes. In some countries people can obtain almost any instrument without any training or formal qualification, in others very high standards of competence are set and people are expected to demonstrate that they can meet these standards before publishers will supply them with materials.

As mobility of test users increases and as the use of test by multinationals across European borders increases, so it has become increasingly important to define some benchmark standards both for test quality and for competence in test use. It is also important for publishers of tests, who operate globally, to be able to set comparable standards for supply in different countries. If they do not, there is the danger of 'grey markets' emerging where poorly qualified test users buy in one country and then use in another.

The EFPA Standing Committee on Tests and Testing (SCTT) has already produced and published standards for the quality of tests. We now needed to address the issue of standards for Test Use.

The development of local national test user qualification systems has been developing in a number of European countries (notable UK, Sweden, Finland and Norway). Those countries and others have expressed a desire for there to be some higher level coordination of these projects to help ensure uniformity of standards between countries.

The work of the current European Working Group sprang from this need and has attempted to address the complex task of defining a meaningful set of standards of competence in test use that can be used as a benchmark for a range of very different national approaches to test user qualification and certification.

What these Standards are about

The standards are about competence in test use. They endeavour to define what people need to know and do, what skills are required and what understanding is required for safe and competent use of a limited range of tests in a variety of occupational settings or contexts (e.g. personnel selection; assessment for management development; guidance for career development; etc). Because testing is such a broad and diverse topic, it is very important for standards processes to make clear the limits of the qualifications that they define. Any qualifications that are developed need to reflect the realities of practice in the field.

The basis chosen for the European Standards was the International Test Commission's International Guidelines on Test Use (Bartram, 2001; ITC, 2001). These have become widely accepted as defining best practice in test use, and have been adopted by a number of psychological associations and translated in many different languages. While these provide a good structure for standards, the ITC Guidelines are not sufficiently specific to provide the basis for qualifications. Consequently, the EWG has been modifying and expanding the ITC Guidelines into a more detailed format. The format adopted is as follows.

The ITC Guidelines have been re-cast into three Units, each containing a number of specific Standards:

Unit 1	**Take responsibility for ethical test use**
Standard: 1.1	**Act in a professional and ethical manner**
Standard: 1.2	**Ensure you have the competence to use tests**
Standard: 1.3	**Take responsibility for their use of tests**
Standard: 1.4	**Ensure that test materials are kept securely**
Standard: 1.5	**Ensure that test results are treated confidentially**
Unit 2	**Follow good practice in the use of tests**
Standard: 2.1	**Evaluate the potential utility of testing in an assessment situation**
Standard: 2.2	**Choose tests appropriate for the situation**
Standard: 2.3	**Give due consideration to issues of fairness in testing**
Standard: 2.4	**Analyse and interpret results appropriately**
Standard: 2.5	**Communicate the results clearly and accurately to relevant others**
Standard: 2.6	**Review the appropriateness of the test and its use**
Unit 3	**Follow good practice in the administration of tests**
Standard: 3.1	**Make necessary preparations for the testing session**
Standard: 3.2	**Administer the tests properly**
Standard: 3.3	**Score the test results accurately**

For each Unit, the general context is defined. For example:

General Occupational Context:
Testing of people for • Recruitment and selection • Organisational or team fit • Identification of development needs • Career advice and guidance Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures)

Administration modes • Individual vs Group • Paper-based, mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration

In addition the knowledge and skills required are outlined. For example:

Knowledge required in this Unit		
Theories, models and principles of:	**Facts, data and information about:**	**Methods, techniques and procedures relating to:**
Ethics and principles of good practice in testing • Informed consent • Privacy and confidentiality Ethics and principles of good practice in client relationships	Psychometric tests Good practice guides: • ITC Guidelines on Test Use • ITC Guidelines on Computer-based testing and testing on the Internet • ITC Guidelines on test adaptation • Local and national good practice guides relating to testing and test use Law and Codes of Practice relating to protection of personal data and intellectual property rights in relation to individual testing Law and Codes of Practice relating to equal opportunities, direct and indirect discrimination	Assessment needs analysis Domain-related: • Recruitment and selection • Organisational or team fit • Identification of development needs • Career advice and guidance

Within each Unit, each standard is defined in terms of what performance is required of the competent test user and the more specific contexts to which the competence relates. For example:

Standard: 1.1	**Act in a professional and ethical manner**	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing people for: • Recruitment and selection • Organisational or team fit • Identification of development needs • Career advice and guidance Stakeholders • Individual • Organisation • Hiring manager • Recruitment consultant Testing methods • Psychometric tests of ability (maximum performance measures) • Psychometric tests of personality, motivation etc (typical performance measures) • Administration modes • Individual vs Group • Paper-based, mechanical, computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Promote and maintain professional and ethical standards.	
b	Have a working understanding of current professional and ethical issues and debates relating to the use of tests.	
c.	Implement an explicit policy on testing and test use.	
d.	Ensure that people who work for or with you adhere to appropriate professional and ethical standards of behaviour.	
e.	Conduct communications and give feedback with due concern for the sensitivities of the test taker and other relevant parties.	
f.	Represent tests and testing in a positive and balanced manner in communications with and through the media.	
g.	Avoid situations in which they may have or be seen to have a vested interest in the outcome of the test, or where the test might damage the relationship with their client.	

In addition there is a specification of the specific knowledge and skills that underpin competence in relation to the Standard.

In addition to the above standards (performance requirements, occupational contexts and specifications of knowledge and skill requirements), the EWG has also produced some example Assessment Guidance. This Guidance is intended to illustrate the level of detail, depth and breadth of the competence defined by the standard, and to provide support for people who are assessing individuals for their competence in test use.

The Standards have the most important role in defining what the EWG has agreed to be the scope and content of qualification in test use. The Assessment Guidance provides practical support for those who might wish to implement a test user certification process based on these standards.

A full set of the standards and the assessment guidance are attached as Appendix A to this report.

3. FUTURE TASKS FOR A JOINT WORKING GROUP

Once the Test User Standards have been accepted by the EAWOP General Assembly (May 2005) and by the EFPA General Assembly (July 2005) as a sound basis for further development, a second phase of the project will be launched.

1. Firstly we should define the principles, rules and procedural practices regarding how the standards could be adapted and applied at the national and local level.
2. Secondly we should establish some principles or minimum criteria regarding how the standards can be used in order to grand certifications for test users.
3. Thirdly the decision making procedures and administrative issues should be considered.

European Test User Standards impose the minimum criteria every national and local adaptation should follow. The questions that need to be taken into consideration and elaborated are:

1. How is the subsidiarity principle to be carried out in practice? What are the decisions that are allowed or must be made at the national level as opposed to those made at the European level? For example, at the national level it may be desired to adopt stricter standards or there may be a different qualification model followed. While the European standard may set minimum criteria for some aspects of local qualifications, it is not necessary for the local qualification to cover exactly the same grounds as the European standards. Local qualifications may be more extensive in some cases, or only meet some of the European requirements in others. A key factor will be the adoption of good practice and ethical guidelines.
2. In every country only those standards that are accepted locally should be considered for accreditation for the European qualification. The issue of how these different national standards and accreditation systems could be compared or managed is a complicated one and similar in principle to the process relating to the European Directive on the recognition of professional qualifications in relation to the establishment of a 'common platform'. This process begins with an inventory that maps qualifications in one country to those in others.
3. How will accreditation be administrated and certification managed for the test users? There is a need to set up some procedural standards for this. Such standards need to cover a host of questions:
    a. Who is eligible to apply the certification?
    b. How can assessment be validated?
    c. For how long is the certification valid?
    d. What evidence may be required on the continued maintenance and development of one's competence as test user?
    e. Under what conditions would a test user forfeit the certification?
    f. Could there be some transition provisions (i.e. 'grand-parenting' arrangements for those who have worked in the field for years could have the certification automatic)?

4. Who is permitted to develop and maintain national standards and grant the certificates? The responsible local body should at least be ruled by an Association that is a member of EFPA or EAWOP. Is there need to regulate this in a more detailed way?

4. IMPLICATIONS FOR EFPA: Administration at the European level

At the European level the most important issue is to establish and agree on the roles of EFPA and EAWOP in decision making and administration.

The core issue is that EFPA and EAWOP have jointly developed and decided on the European Standards. Any change in the Standards shall be accepted by EFPA and EAWOP jointly also in the future. There should be some procedure to evaluate the Standards and their usage on the regular base, too.

Another issue relates to the most convenient way to manage the implementation, accreditation and administration of the standards in practice at the European level. One option is that EAWOP takes the prime responsibility of this. This could make the management more simple and flexible. EAWOP contains the relevant subject-matter experts in Work and Organizational psychology. Also, as an organisation, EAWOP has individual members. Thus EAWOP has already developed administrative mechanism to deal with registers of individuals and gather payments from them if these kinds of solutions are needed.

Nevertheless there are a number of open practical questions to be solved:

1. How to establish and run a body that evaluates the national standards and practices and accredits them?
2. How the European body or committee would work and makes decisions?
3. What mechanism there is for someone to make a complaint or appeal against the body's decisions?
4. How to handle the complaints of the individuals rejected by the national bodies?
5. What kind of register is needed and how to manage it?
6. How costs and expenses are covered? There could be some accreditation payment for the national organisations. There could also be some payments for individuals who want to be registered at the European level.

In most cases, the above issues have already been addressed in a broader forum in relation to the European Diploma in Psychology. We would intend to build on that foundation and take the EuroPsy as a useful starting point. We would also seek to draw practical solutions for Test User certification from the ongoing 'experimental garden' EuroPsy projects.

This cross-fertilisation would be facilitated by the fact that the first author is a member of the EuroPsy Leonardo project group.

A common EFPA and EAWOP position for the future

We propose that

- A small joint EFPA-EAWOP implementation group is established to explore issues relating to the above. This new group would produce a detailed implementation plan in time for the EFPA and EAWOP General Assemblies in 2007.
- We adopt processes similar to those suggested for the European Diploma in Psychology, whereby local arrangements can be accredited if they meet the requirements for the Diploma, and so much of the administrative load is devolved to local level.

5. SPECIFIC RECOMMENDATIONS

The EFPA-EAWOP EWG on Test User Qualification makes the following recommendations to the EFPA and EAWOP General Assemblies.

It is recommended that:

1. The standards (see Annex) are recognised by the General Assemblies as an acceptable basis for the further development of European Standards for Test Use in the field of Work and Organizational assessment.
2. Through its Standing Committee on Tests and Testing, EFPA establishes procedures to support and maintain these standards.
3. That EFPA and EAWOP, through their Executive Committees, establish a small joint board responsible for reviewing progress, and dealing with policy issues.
4. EAWOP takes the lead role in implementing a European process for accrediting local national qualifications in test use based on the Standards and for establishing a European Register of Test Users.
5. A small joint EFPA-EAWOP implementation group is established to explore issues relating to the above.
    1. Complete the standards development and definition work in the light of the feedback from the consultation
    2. Develop a set of rules and regulations regarding European accreditation and European certification, using the European Diploma in Psychology regulations as a model.
    3. Collaborate with current certification providers and pilot projects launched at the national levels.
    4. Prepare a proposal for how the Standards and the accreditation is administrated and managed at the European level, including a definition of the responsible partners and bodies.
    5. Produce detailed recommendations on implementation for the EFPA and EAWOP Executive Committees in 2007 with final proposals for the standards and their implementation.

References

Bartram, D. (2001). The development of international guidelines on test use: The International Test Commission project. *International Journal of Testing, 1,* 33-55.

ITC (2001). International Guidelines for Test Use. *International Journal of Testing, 1,* 93-114.

Annexes

Membership of the EFPA-EAWOP European Working Group

Name	Country	Representing	Role
Dave Bartram	UK	EFPA SCTT	EWG Convenor
Johanne Bratbo	Denmark	EFPA SCTT	Executive Council Liaison
Kathia Glabeke	Belgium	EFPA SCTT	
Daniel Kozeny	Czech Republic	EFPA SCTT	
Miloslav Solc	Czech Republic	EFPA SCTT	
Helle Milton	Denmark	EFPA SCTT	
Harald Ackerschott	Germany	EFPA SCTT	
Gerd Reimann	Germany	EFPA SCTT	
Boele de Raad	Netherlands	EFPA SCTT	
Wouter Schoonman	Netherlands	EFPA SCTT	
Kjetil Soren Sundet	Norway	EFPA SCTT	
Jose Muniz	Spain	EFPA SCTT	
Christina Stromer-Wilson	Sweden	EFPA SCTT	
Orjan Frans	Sweden	EFPA SCTT	
Pat Lindley	UK	EFPA SCTT	
Henry Honkanen	Finland	EAWOP	EWG Co-Convenor
Elmar Lamerskitten	Germany	EAWOP	
Olaf Ringelband	Germany	EAWOP	
Vicente Gonzalez-Roma	Spain	EAWOP	
Helen Baron	UK	EAWOP	
Hayley Oldale	UK	BPS WG	Administrator
Jim Boyle	UK	BPS WG	
Linda Marshall	UK	BPS WG	
Mike Wang	UK	BPS WG	
Sue Waters	UK	BPS WG	

Note: EFPA SCTT = EFPA Standing Committee on Tests and Testing.
BPS WG: Members of the British Psychological Society's Working Group on the review of test user standards.

Attached as separate documents.

Current versions of:

- EWG2: The Consultation questions
- EWG3: The Test Use Standards
- EWG4: The glossary to the standards
- EWG5: Draft Assessment Guidance
- EWG6: Aggregated feedback from the consultation

Dave Bartram
Dave.Bartram@shlgroup.com

EWG2

EFPA Standards for Test User Qualification

Consultation for the draft European Standards for Competence in the Use of Tests in Work and Organizational Settings (the European Standards)

As the title specifies, the draft European Standards for Competence in the Use of Tests in Work and Organizational Settings (the European Standards) have been developed to provide a common European specification of the competences expected of people using tests in work and organizational settings. These standards are intended to help those who have or who are intending to develop qualifications in this area to adopt a common standard or benchmark against which they can set their local qualifications. They are not intended for use in other settings, such as education or health care.

It is envisaged that the European Standard could become the basis for an accreditation process that could provide local qualifications with European recognition.

The standards have been developed by a joint EFPA and EAWOP Working Group and have used the International Test Commission's Guidelines on Test Use (for details, download a copy from www.intestcom.org). as the basis for the structure and content. They have also been developed to be consistent with standards that currently form the basis for test user qualifications in the UK, Finland, Sweden and Norway. The development of these standards has been carried out through a series of European workshops and consultation procedures under the auspices of the relevant local national psychological associations and of EFPA and EAWOP. We have now reached a point in this work where more wide-reaching consultation is needed. We invite you to provide this input.

For more details of the remit of the EFPA-EAWOP Working Group charged with doing this work, see Bartram (2004) [European Psychologist].

The main documents on which this consultation is based include:

1. The Standards themselves. This document specifies the performance required of a competent test user and defines the contexts within which that person would operate. It also sets out specifications of the knowledge and skills that underpin competence. The first part of this document describes the structure with examples of the content. This document also contains a glossary of definitions of the terms used.
2. A Glossary that defines the terms used in the Standards.
3. An example of an Assessment Guidance document based on the Standards to show how the standards can be turned into qualifications. This has been created by mapping the UK's British Psychological Society's Level A and Level B qualification requirements to the Standards. This assessment guidance is intended to guide those who would be judging someone's competence in relation to the standards. Each country may define this level of detail differently, to reflect local differences in practice. This is not an issue so long as the definition of competence meets the standards required.

We are asking people to focus their feedback and comments on Document 1: The Standards. The Assessment Guidance is provided simply to illustrate how the standards might be expanded to define qualification requirements.

Qualification systems and the standards

The full set of Standards (Document 1) defines a comprehensive qualification in test use. In practice, people use tests in relatively restricted setting or contexts that would not require attainment of the full range of competencies specified by the standards. In the UK for example, there are qualifications that recognise competence in test administration, competence in test use relating to measurement of ability, and competence in test use with specific personality instruments. In addition there is a more comprehensive qualification that recognises those who have specialised in testing and assessment in this field, rather than those who use it as one part of their everyday function.

We regard it as important to provide flexibility in how qualifications are constructed from the standards in order to reflect local practice and custom, while retaining the notion of a comprehensive qualification covering all the standards and content. This latter qualification sets a 'gold standard'.

While individual countries are free to establish their own qualifications structures, we would like to know whether you believe it would be of value to provide a mechanism for providing some recognition of these local qualifications as providing partial fulfilment of or credit towards the full European standard. This could be a very strong incentive for people to seek full qualification.

Consultation Questionnaire

Please complete as much as you can of the following questionnaire once you have reviewed the relevant documents.

Please give or email the completed questionnaire to your EFPA-EAWOP Working Group representative or email your completed response to Prof Dave Bartram, the convenor of the EFPA-EAWOP Working Group: Dave.Bartram@shlgroup.com

Your details

Q1

Name:	
Contact details (email):	
Contact details (phone):	

Q2: I am responding:

	On behalf of the following European organization, commission or association: [please specify]
	On behalf of the following National organization, commission or association: [please specify]
	On behalf of the following company, organization or group: [please specify,]
	On my own behalf.

Q3: If responding on behalf of a specific committee or group within an organisation, please specify:

I/We have the following involvement in testing [tick as many as apply].

	Professional association of psychologists
	Professional association of non-psychologists who make use of tests
	Trainer or training organisation for test users
	Test user or organisation of test users
	Test taker, group or organisation representing test takers
	Developer of tests
	Publisher/distributor of tests
	Other [please specify]

Q4: Would you like your contact details to be added to our mailing list so that you can be notified of further developments in relation to this project?

	YES
	NO

Consultation Questions

Structure of the standards:

	Very unclear	Unclear	Adequate	Clear	Very clear
Q5: How clear is the way in which the Standards have been structured?					
If unclear, please suggest how might this be improved?					

Content of the standards

	Much too little detail	Too little detail	About right	Too much detail	Much too much detail
Q6: Is the amount of detail provided about right?					
Are there any specific areas where either more or less detail is needed?					

	Very poorly	Poorly	About right	Well	Very well
Q7: In general, how well does the draft Standards cover the areas of competence required of test users in work and organizational settings?					

Q8: Are there any specific aspects of competence that are missing?

	YES
	NO
If YES, can you define what they are?	

Q9: Are any aspects of the competences that are specified not necessary?

	YES
	NO
If YES, can you say which ones?	

Qualifications related to the standards

	Not all valuable	Not valuable	In between	Valuable	Very valuable
Q10: How valuable do you think test users would perceive a European accredited qualification in test use to be?					

Q11: Do you think that test user qualifications in your country should be consistent with this European Standard?

	YES
	NO
If NO, can you can you say why not?	

General comments related to the standards

Q12: If you have any specific comments to make on the content of the draft Standards, please provide those here.	

If you know of any individuals of groups you think should be consulted about these Standards, please copy the documents to them or provide the contact details with your response.

Thank you for taking the time to comment on the propose standards. Your input will be a valuable contribution to the future of this project.

EWG3

Document 1

European Test User Standards

for test use in Work and Organizational settings

CONSULTATION DRAFT

Prepared by

the European Federation of Psychologists' Associations

and

the European Association of Work and Organizational Psychologists

Version 1.8. 25 March 2005

Why Test User Standards are needed?

Within the work and organizational field, there is increasing interest in Europe in the development of test user standards and test standards. This interest arises not only from the psychology profession but also from allied professions, like human resources management.

At present there is little if any uniformity across Europe in terms of test user qualification standards or processes. In some countries people can obtain almost any instrument without any training or formal qualification, in others very high standards of competence are set and people are expected to demonstrate that they can meet these standards before publishers will supply them with materials.

As mobility of test users increases and as the use of test by multinationals across European borders increases, so it has become increasingly important to define some benchmark standards both for test quality and for competence in test use. It is also important for publishers of tests, who operate globally, to be able to set comparable standards for supply in different countries. If they do not, there is the danger of 'grey markets' emerging where poorly qualified test users buy in one country and then use in another.

The EFPA Standing Committee on Tests and Testing (SCTT) has already produced and published standards for the quality of tests. The present Standards address the issue of standards for Test Use.

The development of local national test user qualification systems has been developing in a number of European countries (notably UK, Sweden, Finland and Norway). Those countries and others have expressed a desire for there to be some higher level coordination of these projects to help ensure uniformity of standards between countries.

The work of the joint EFPA-EAWOP European Working Group (EWG) sprang from this need and has attempted to address the complex task of defining a meaningful set of standards of competence in test use that can be used as a benchmark for a range of very different national approaches to test user qualification and certification.

What these Standards are about?

The standards are about competence in test use. They endeavour to define what people need to know and do, what skills are required and what understanding is required for safe and competent use of a limited range of tests in a variety of occupational settings or contexts (e.g. personnel selection; assessment for management development; guidance for career development; etc). Because testing is such a broad and diverse topic, it is very important for standards processes to make clear the limits of the qualifications that they define. Any qualifications that might be based on these standards need to reflect the realities of practice in the field.

The basis chosen for the Standards was the International Test Commission's (ITC) International Guidelines on Test Use (Bartram, 2001; ITC, 2001 [check ref]). These have become widely accepted as defining best practice in test use, and have been adopted by a number of psychological associations and translated into many different languages. While these provide a good structure for standards, the ITC Guidelines are not sufficiently specific to provide the basis for qualifications. Consequently, the EWG has been modifying and expanding the ITC Guidelines into a more detailed format. The format adopted is as follows.

The ITC Guidelines have been re-cast into three Units, each containing a number of specific Standards:

Unit 1	**Take responsibility for ethical test use**
Standard: 1.1	**Act in a professional and ethical manner**
Standard: 1.2	**Ensure you have the competence to use tests**
Standard: 1.3	**Take responsibility for their use of tests**
Standard: 1.4	**Ensure that test materials are kept securely**
Standard: 1.5	**Ensure that test results are treated confidentially**
Unit 2	**Follow good practice in the use of tests**
Standard: 2.1	**Evaluate the potential utility of testing in an assessment situation**
Standard: 2.2	**Choose tests appropriate for the situation**
Standard: 2.3	**Give due consideration to issues of fairness in testing**
Standard: 2.4	**Analyse and interpret results appropriately**
Standard: 2.5	**Communicate the results clearly and accurately to relevant others**
Standard: 2.6	**Review the appropriateness of the test and its use**
Unit 3	**Follow good practice in the administration of tests**
Standard: 3.1	**Make necessary preparations for the testing session**
Standard: 3.2	**Administer the tests properly**
Standard: 3.3	**Score the test results accurately**

For each Unit, the general occupational context is defined.

General Occupational Context:
Testing of people for • Recruitment and selection • Organisational or team fit • Identification of development needs • Career advice and guidance Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration

In addition the knowledge and skills required are outlined. For example:

Knowledge required in this Unit		
Theories, models and principles of:	**Facts, data and information about:**	**Methods, techniques and procedures relating to:**
Ethics and principles of good practice in testing • Informed consent • Privacy and confidentiality Ethics and principles of good practice in client relationships	Psychological tests Good practice guides: • ITC Guidelines on Test Use • ITC Guidelines on Computer-based testing and testing on the Internet • ITC Guidelines on test adaptation	Assessment needs analysis Domain-related: • Recruitment and selection • Organisational or team fit • Identification of development needs • Career advice and guidance

	• Local and national good practice guides relating to testing and test use Etc…	

Within each Unit, each standard is defined in terms of what performance is required of the competent test user and the more specific contexts to which the competence relates. For example:

Standard: 1.1	**Act in a professional and ethical manner**	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing people for: • Recruitment and selection • Organisational or team fit • Identification of development needs • Career advice and guidance Stakeholders • Individuals • Organizations • Hiring managers • Recruitment consultants • HR specialists Testing methods • Psychometric tests of ability (maximum performance measures) • Psychometric tests of personality, motivation etc (typical performance measures) • Administration modes • Individual vs Group • Paper-based, mechanical, computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Promote and maintain professional and ethical standards.	
b	Have a working understanding of current professional and ethical issues and debates relating to the use of tests.	
c.	Implement an explicit policy on testing and test use.	
d.	Ensure that people who work for or with you adhere to appropriate professional and ethical standards of behaviour.	
e.	Conduct communications and give feedback with due concern for the sensitivities of the test taker and other relevant parties.	
f.	Represent tests and testing in a positive and balanced manner in communications with and through the media.	
g.	Avoid situations in which they may have or be seen to have a vested interest in the outcome of the test, or where the test might damage the relationship with their client.	

In addition there is a specification of the knowledge and skills that underpin competence in relation to the Standard.

In addition to the Standards (performance requirements, occupational contexts and specifications of knowledge and skill requirements), the EWG has also produced some example Assessment Guidance (contained in a separate document). The Assessment Guidance is intended to illustrate the level of detail, depth and breadth of the competence defined by the standard, and to provide support for people who are assessing individuals for their competence in test use.

The Standards have the most important role in defining what the EWG has agreed to be the scope and content of qualification in test use. The Assessment Guidance provides practical support for those who might wish to implement a test user certification process based on these standards.

A GLOSSARY has been produced to clarify the definition of terms used in the Standards. The Glossary is provided as a separate document.

The Standards

Unit 1	Take responsibility for ethical test use
Standard: 1.1	Act in a professional and ethical manner
Standard: 1.2	Ensure you have the competence to use tests
Standard: 1.3	Take responsibility for their use of tests
Standard: 1.4	Ensure that test materials are kept securely
Standard: 1.5	Ensure that test results are treated confidentially

Unit 2	Follow good practice in the use of tests
Standard: 2.1	Evaluate the potential utility of testing in an assessment situation
Standard: 2.2	Choose tests appropriate for the situation
Standard: 2.3	Give due consideration to issues of fairness in testing
Standard: 2.4	Analyse and interpret results appropriately
Standard: 2.5	Communicate the results clearly and accurately to relevant others
Standard: 2.6	Review the appropriateness of the test and its use

Unit 3	Follow good practice in the administration of tests
Standard: 3.1	Make necessary preparations for the testing session
Standard: 3.2	Administer the tests properly
Standard: 3.3	Score the test results accurately

Unit 1	Take responsibility for ethical test use
Standard: 1.1	Act in a professional and ethical manner
Standard: 1.2	Ensure you have the competence to use tests
Standard: 1.3	Take responsibility for their use of tests
Standard: 1.4	Ensure that test materials are kept securely
Standard: 1.5	Ensure that test results are treated confidentially

What is this unit about?	This unit is about taking responsibility for acting in an ethical manner
Who is this unit for?	This unit should be appropriate for all test users involved in or advising on individual assessments which include or might include the use of psychological tests.

Occupational Context:

Testing of people for
- Recruitment and selection
- Organizational or team fit
- Identification of development needs
- Career advice and guidance
- Other organizational purposes

Testing of groups for
- Team development
- Competencies audit
- Other organizational purposes

Testing methods
- Psychological tests of ability (maximum performance measures)
- Psychological tests of personality, motivation etc (typical performance measures)

Administration modes
- Individual vs Group
- Paper-based, Mechanical, Computer-based
- Online (internet) vs offline testing
- Open, Controlled, Supervised or Managed administration

Knowledge required in this Unit

Theories, models and principles of:	Facts, data and information about:	Methods, techniques and procedures relating to:
Ethics and principles of good practice in testing including: • Informed consent • Privacy and confidentiality Ethics and principles of good practice in client relationships	Psychological tests Good practice guides: • ITC Guidelines on Test Use • ITC Guidelines on Computer-based testing and testing on the Internet • ITC Guidelines on test adaptation • Local and national good practice guides relating to testing and test use Law and Codes of Practice relating to protection of personal data and intellectual property rights in relation to individual testing Law and Codes of Practice relating to equal opportunities, direct and indirect discrimination, employment law.	Assessment needs analysis Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance

Standard: 1.1	**Act in a professional and ethical manner**	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance Testing of groups for • Team development • Competencies audit • Other organizational purposes Stakeholders • Individuals • Organizations • Hiring managers • Recruitment consultants • HR specialists Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Promote and maintain professional and ethical standards.	
b	Implement an explicit policy on testing and test use.	
c	The people who work for or with you adhere to appropriate professional and ethical standards of behaviour.	
d	Conduct communications and give feedback with due concern for the sensitivities of the test taker and other relevant parties.	
e	Represent tests and testing in a positive and realistic manner in communications with and through the media.	
f	Avoid situations in which you may have or be seen to have a vested interest in the outcome of the test, or where the test might damage the relationship with your client.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Professional and ethical standards in testing Professional issues and debates relating to use of tests in occupational settings Testing and test use policies	Ability to use appropriate levels and methods of communication, taking account of the needs, abilities and sensitivities of test takers and other relevant parties. Ability to present tests and testing in a positive and balanced manner.

Standard: 1.2	**Ensure you have the competence to use tests**	
The Occupational Standard		
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance Testing of groups for • Team development • Competencies audit • Other organizational purposes Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Work within the limits of scientific principle and empirical evidence.	
b	Have sufficient knowledge of test theory and psychometrics for the use you will make of tests.	
c	Know the limits of your own competence and operate within those limits in terms of contexts of testing, testing methods used and modes of administration.	
d	Keep up with relevant changes and advances relating to the testing methods you use and with advances in the tests you use.	
e	Set and maintain high personal standards of competence.	
f	Keep up with relevant changes and developments in legislation and policy, which may impact on your use of tests.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Scientific method. Knowledge of own competence and qualifications Up-to-date knowledge of 1. testing methods and tests in general use in the occupational field, changes in testing 2. specific tests one uses 3. relevant developments in legislation and policy	Research and Study skills (for maintaining knowledge) Learning from experience Self-reflection Self-awareness Search skills (for locating information about tests and testing)

Standard: 1.3	**Take responsibility for their use of tests**	
The Occupational Standard		
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance Testing of groups for • Team development • Competencies audit • Other organizational purposes Participants • Test taker • Organizational representatives • Hiring manager • Recruitment consultant • Career advisor Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Only offer testing services, modes of administration and testing methods for which you are qualified.	
b	Accept responsibility for the choice of tests used, and for the recommendations made.	
c	Provide clear and adequate information to participants in the testing process about the ethical principles and legal regulations governing psychological testing.	
d	Ensure other parties are informed of their duties of responsibility and accountability regarding legal, contractual and ethical codes	
e	Accept the duties of responsibility and accountability regarding legal, contractual and ethical codes.	
f	Ensure that there is informed consent and that the nature of the contract between test-taker (or responsible person) and tester is clear and understood.	
9	Be alert to any unintended consequences of test use.	
h	Try to avoid doing harm or causing distress to those involved in the testing process.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Ethical principles and legal regulations relating to psychological testing. Psychological contract between test-taker and other parties involved in the test process. Legal, contractual and ethical codes.	Production of test-taker contracts Communication skills Ability to maintain and manage appropriate professional relationships with participants

Standard: 1.4	**Ensure that test materials are kept securely**	
The Occupational Standard		
	Performance required	**Occupational Context**
	You must ensure that you:	Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Safeguard access to test materials e.g. store test materials securely and control access to them.	
b	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.	
c	Protect the security of the test e.g. by not coaching individuals on actual test materials or other practice materials that might unfairly influence their test performance.	
d	Ensure that test techniques are not described publicly or in other ways placed in the public domain in such a way that their usefulness is impaired	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Copyright law and related restrictions on distributions of materials Risks associated with different modes of administration and storage methods	Management of testing materials Management of data, filing and data storage

Standard: 1.5	**Ensure that test results are treated confidentially**	
The Occupational Standard		
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance Testing of groups for • Team development • Competencies audit • Other organizational purposes Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Filing systems • Computerised • Paper-based Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Specify who will have access to results and define levels of confidentiality.	
b	Explain levels of confidentiality to individuals before tests are administered.	
c	Limit access to results to those with a right to know.	
d	Obtain the relevant consents before releasing results to others.	
e	Protect data kept on file systems so that only those who have a right of access can obtain them.	
f	Establish clear guidelines as to how long identifiable personal test data are to be kept on file.	
g	Remove names and other personal identifiers from databases of results that are archived for research use, development of norms or other statistical purposes.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Data Protection legislation Good practice codes relating to the management of personal data Procedures for ensuring security of paper and computer-based personal data records. Anonymisation of data records.	Communication skills Management of data, filing and data storage

Unit 2	Follow good practice in the use of tests
Standard: 2.1	Evaluate the potential utility of testing in an assessment situation
Standard: 2.2	Choose tests appropriate for the situation
Standard: 2.3	Give due consideration to issues of fairness in testing
Standard: 2.4	Analyse and interpret results appropriately
Standard: 2.5	Communicate the results clearly and accurately to relevant others
Standard: 2.6	Review the appropriateness of the test and its use

What is this unit about?	This unit is about following good practice in choosing, using and interpreting psychological tests
Who is this unit for?	This unit should be appropriate for all test users involved in or advising on individual assessments which include or might include the use of psychological tests.

Occupational Context:

Testing of people for:
- Recruitment and selection
- Organizational or team fit
- Identification of development needs
- Career advice and guidance
- Other organizational purposes

Testing of groups for:
- Team development
- Competencies audit
- Other organizational purposes

Testing methods
- Psychological tests of ability (maximum performance measures)
- Psychological tests of personality, motivation etc (typical performance measures)

Administration modes
- Individual vs Group
- Paper-based, Mechanical, Computer-based
- Online (internet) vs offline testing

Open, Controlled, Supervised or Managed administration

Knowledge required in this Unit

Theories, models and principles of:	Facts, data and information about:	Methods, techniques and procedures relating to:
Test-related theories and models: • Classical Test theory • Item Response Theory • Fairness and bias in testing • Reliability and Validity • Generalisability theory Psychological theories and models of • Cognitive ability • Personality • Motivation • Attitudes, beliefs and values • Other constructs used in testing in this occupational context Theories and models relating to testing in the workplace: • Workplace competencies • Performance management • Personal development • Career choice • Selection and recruitment Ethics and principles of good practice in testing including: • Informed consent Privacy and confidentiality	Psychological tests Good practice guides: • ITC Guidelines on Test Use • ITC Guidelines on Computer-based testing and testing on the Internet • ITC Guidelines on test adaptation • Local and national good practice guides relating to testing and test use Law and Codes of Practice relating to protection of personal data and intellectual property rights in relation to individual testing Law and Codes of Practice relating to equal opportunities, fairness. Employment law	Assessment needs analysis Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes

Standard: 2.1	**Evaluate the potential utility of testing in an assessment situation**	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes Stakeholders • Individuals • Organizations • Hiring managers • Recruitment consultants • HR specialists Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Base decisions about testing on a thorough analysis of the client's requirements, ensuring that you fully understand the assessment need.	
b		
c		
d	Establish that the constructs being measured are relevant for the assessment need Assess advantages and disadvantages of using tests instead of or with other sources of information or alternative assessment strategies	
e	Produce a reasoned justification for the use of tests in relation to: • assessment methods • modes of administration, • stakeholders • and reasons for testing.	
f	Conduct a risk assessment with regards to assessment method security, legal issues and protection of intellectual property in relation to: • testing methods • testing conditions • modes of administration, • context, • stakeholders.	
g	Consider how other sources of information might be used together with or instead of the tests. Can explain to the relevant stakeholder the utility associated with the use of tests and other methods of assessment.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Job and task analysis Competency profiling Legal and best practice requirements relating to confidentiality and security of personal data Costs, benefits and risks associated with different modes of testing Intellectual property rights and copyright law Guidelines of relevant national and international professional bodies Utility theory	Job and task analysis methods Competency profiling methods Relating job descriptions to person specifications Risk assessment Cost-benefit analysis

Standard: 2.2	Choose tests appropriate for the situation	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes Criteria of appropriateness • Scope or coverage • Reliability • Validity • Fairness • Acceptability to stakeholders • Practicality in terms of admin, cost etc Stakeholders • Individuals • Organizations • Hiring managers • Recruitment consultants • HR specialists Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	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.	
b	Consider which tests, if any, are appropriate for the assessment contexts	
c	Choose tests based on an evaluation of the technical and user documentation and other available information to ensure that the test's scope, reliability, validity, standardisation, fairness, accessibility, language or dialect version, practicality and acceptability are appropriate to the identified assessment need.	
d	Agree the choice of tests, administration conditions and administration modes with relevant stakeholders.	
e	Respond to requests from test takers by providing sufficient information to allow them to understand why the test was chosen.	
f	Agree applications of results and confidentiality and security of the personal data with relevant stakeholders.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Relevant testing methods and instruments Validity and reliability Measurement error Classical Test Theory Item Response Theory Norms and standardisation Differential item functioning and test bias Psychometric properties of testing instruments. Legal and practice requirements relating to confidentiality and security of personal data	Assessment needs analysis Negotiating and agreeing testing procedures with stakeholders. Integration of testing into broader assessment procedures Use of assessment matrix to map constructs being assessed against methods of assessment

Standard: 2.3	**Give due consideration to issues of fairness in testing**	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes Various groups differing in terms of • gender, • cultural background, • education, • ethnic origin, • religion • age • sexual orientation • language • any disability Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration Appropriate arrangements for: • Hearing, visual or motor impairments • Learning disabilities, dyslexia
a	Avoid the use of stereotypes in test selection, administration and interpretation	
b	Evaluate any potential for bias in the test given the testing context and conditions, for the various groups that will be tested.	
c	Check: i. that the modes of administration are appropriate ii. that the constructs being assessed are meaningful in each of the groups represented iii. that evidence is available on possible group differences in performance on the test. iv. that there is validity evidence to support the intended use of the test in the various groups and that the effects of any group differences not relevant to the main purpose of assessment (e.g., differences in motivation to answer, or reading ability) are minimised. v. that the context, culture and language of the test is appropriate.	
d	Interpret guidelines relating to the fair use of tests in the context of local policy and legislation.	
e	Make appropriate arrangements for test takers with disabilities.	
f	Ensure that test administrators are aware of issues of fairness and take them into account before during and after administration.	
g	Check that administrators: i. implement accommodations for test takers with disabilities appropriately ii. can communicate proficiently with the test taker in an appropriate language. iii. Observe tests takers and identify any difficulties they may have with the test process iv. Document any administration problems or deviations from standard procedures.	
h	Analyse and interpret scores with due regard for fairness and equity issues.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Relevant testing methods and instruments Validity and reliability Measurement error Classical Test Theory Item Response Theory Norms and standardisation Differential item functioning and test bias Psychometric properties of testing instruments. Legal and practice requirements relating to equal opportunities and disability ITC test adaptation guidelines Impact of various disabilities on test taking performance. Best practice on making allowances for various forms of impairment of disability.	

Standard: 2.4	**Analyse and interpret results appropriately**	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes Test taker information • age, • gender, • schooling, • culture • other factors Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Other sources of information • Assessment and development centre exercises • Structured interviews • Biodata and related inventories • Simulations and behavioural (e.g. work) samples Administration modes • Individual vs Group • Paper-based vs Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Produce an interpretation of the results that takes account of available information about the test taker with due regard to the testing context, conditions, methods and modes; the technical qualities of the test; and the assessment need.	
b	Choose suitable norm or reference groups as appropriate for the context in which testing took place and the testing methods used	
C	Choose scale types and derived scores relevant to the intended use of the test scores e.g. stens, percentiles. Consider reasonable alternative interpretations	
d	Produce an interpretation of the results that takes account of available information from other sources or other methods of assessments.	
e	Give due consideration to the available evidence of the technical properties of the test for members of the test takers' relevant demographic groups.	
f	Take account of the possible impact of prior experience on the test	
g	Avoid over-generalising the results of one test	
h	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.	
i	Take account of any individual or group variations from standard procedures in test administration.	
j	Use computer generated reports appropriately in interpretation and with due regard for their technical qualities.	
k		

The Learning Specification	
Essential Knowledge	**Essential Skills**
General knowledge of the psychometric properties of testing instruments. Knowledge of specific tests used: theoretical or conceptual basis, technical documentation and guidance on the use and interpretation of the scale scores. Knowledge of different scale types (normative, ipsative) and item response formats (forced choice, open, rating) and their effects on scale score interpretation. Knowledge of scales used, the characteristics of the norm or comparison groups, and the limitations of the scores Norms and standardisation Different types of standard score scales and methods of converting between them.	Test scoring Conversion of test raw scores to standard scores Bringing together and documenting of tests scores and other assessment data for reports. Computation, where appropriate, of composite scores using standard formulae and equations. Use of norm tables General test interpretation skills Interpretation of ipsative instruments Context related interpretation skills Instrument specific interpretation skills Relating test scores to other information that relates to measurement of the same constructs (e.g. interview ratings, assessment centre ratings etc)

Standard: 2.5	Communicate the results clearly and accurately to relevant others	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes Clients • Individuals • Organizations • Hiring managers • Recruitment consultants Stakeholders • The test taker • The client • Relevant third parties Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Identify relevant others who may legitimately receive test results.	
b	Discuss results with test takers and relevant others in a constructive and supportive manner.	
c	Agree final outcomes and reports with relevant others	
d	Communicate results to agreed recipients.	
e	Provide written or oral reports for relevant interested parties, with the informed consent of the test takers, or their legal representatives.	
f	Use a form and structure for a report that is appropriate to the context of the assessment.	
g	Include within written reports a clear summary, and when relevant, specific recommendations.	
h	Ensure that the technical and linguistic levels of any reports are appropriate for the level of understanding of the recipients.	
i	Explain how the importance of the test results should be weighted in relation to other information about the people being assessed.	
j	Make clear that the test data represent just one source of information and should always be considered in conjunction with other information.	
k	When appropriate provide decision-makers with information on how results may be used to inform their decision	

The Learning Specification	
Essential Knowledge	**Essential Skills**
General knowledge of the psychometric properties of testing instruments. Knowledge of specific tests used Validity and reliability Measurement error Classical Test Theory Item Response Theory Norms and standardisation Differential item functioning and test bias	Report writing Providing feedback Integration of test results with other assessment outcomes for reports. Give appropriate weight to the results and communicate this effectively to the relevant stakeholders.

Standard: 2.6	**Review the appropriateness of the test and its use**	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes
a	Monitor and periodically review change over time in the populations of individuals being tested and any criterion measures being used.	
b	Monitor test data for evidence of adverse impact and re-evaluate the legal and equity issues arising in using the test and check that best practice is being followed.	
c	Validate test for the use to which it is being put, or participate in formal validation studies where possible and practical.	
d	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.	
e	Re-evaluate the use of a test if changes are made to its form, content, or mode of administration.	
f	Re-evaluate the evidence of validity if the purpose for which a test is being used is changed.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
General knowledge of the psychometric properties of testing instruments. Knowledge of specific tests used Validity and reliability Norms and standardisation	Design of validation studies

Unit 3	Follow good practice in the administration of tests
Standard: 3.1	Make necessary preparations for the testing session
Standard: 3.2	Administer the tests properly
Standard: 3.3	Score the test results accurately

What is this unit about?	This unit is about following good practice in administering psychological tests
Who is this unit for?	This unit should be appropriate for all those involved in the administration of tests to individuals or groups.

Occupational Context:

Testing of people for:
- Recruitment and selection
- Organizational or team fit
- Identification of development needs
- Career advice and guidance
- Other organizational purposes

Testing of groups for:
- Team development
- Competencies audit
- Other organizational purposes

Testing methods
- Psychological tests of ability (maximum performance measures)
- Psychological tests of personality, motivation etc (typical performance measures)

Administration modes
- Individual vs Group
- Paper-based, Mechanical, Computer-based
- Online (internet) vs offline testing
- Open, Controlled, Supervised or Managed administration

Knowledge required in this Unit

Theories, models and principles of:	Facts, data and information about:	Methods, techniques and procedures relating to:
Ethics and principles of good practice in test administration • Informed consent • Privacy and confidentiality	Psychological tests Good practice guides relating to test administration: • ITC Guidelines on Test Use • ITC Guidelines on Computer-based testing and testing on the Internet • Local and national good practice guides relating to testing and test use Law and Codes of Practice relating to protection of personal data and intellectual property rights in relation to testing Law and Codes of Practice relating to equal opportunities, direct and indirect discrimination, and employment law.	Assessment needs analysis Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes

Standard: 3.1	Make necessary preparations for the testing session	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Provide relevant parties in a timely manner with clear information concerning the purpose of testing, how the results will be used and data management arrangements.	
b	Provide guidance on ways in which test takers might best prepare for the test session, and the procedures to be followed.	
c	Check that the test is appropriate for the linguistic or dialectic group of the test taker.	
d	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.	
e	Explain clearly to test takers their rights and responsibilities.	
f	Gain the explicit informed consent of test takers or their representatives before any testing is done.	
g	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.	
h	Make the necessary practical arrangements for the test sessions.	
i	Ensure test centres or locations are available and suitably equipped.	
j	Ensure that staff who will be involved in the administration are competent.	
k	Inform the test taker of the testing context and conditions	
l	Prepare the testing methods, equipment and materials	
m	Make arrangements for the testing of people with disabilities	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Relevant testing methods and instruments	Management of testing materials General administration

Standard: 3.2	Administer the tests properly	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advic e and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration Technical problems and issues • Computer system crashes • Loss of Internet connection Personal problems and issues • Illness or sickness • Anxiety • Session interruptions or disruptions
a	Welcome test-takers and brief them in a positive fashion, and act to reduce their anxiety.	
b	Provide appropriate assistance to test takers who show signs of undue distress or anxiety.	
c	Carry out administration procedures as specified in test manuals.	
d	Deal appropriately with any questions, technical or personal problems or issues arising during the testing session.	
e	Observe and record deviations from test procedures.	
f	Ensure the security and safety of testing materials and that all materials are accounted for at the end of each testing session	
g	Adhere strictly to the directions and instructions specified in test manuals while making reasonable accommodations for persons with disabilities.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
The possible impact of assessment on test takers The possible impact of mode of administration on the quality of test results The possible impact of anxiety on test performance.	General skills associated with managing issues and problems that might arise in testing sessions Test administration process skills relating to different types of tests and conditions of administration (group, individual; interactive non-interactive) Management of testing materials

Standard: 3.3	Score test results accurately	
	Performance required	**Occupational Context**
	You must ensure that you:	Testing of people for: • Recruitment and selection • Organizational or team fit • Identification of development needs • Career advice and guidance • Other organizational purposes Testing of groups for: • Team development • Competencies audit • Other organizational purposes Testing methods • Psychological tests of ability (maximum performance measures) • Psychological tests of personality, motivation etc (typical performance measures) Administration modes • Individual vs Group • Paper-based, Mechanical, Computer-based • Online (internet) vs offline testing • Open, Controlled, Supervised or Managed administration
a	Score test results according to standardised procedures where scoring is not automated.	
b	Generate normed and derived scores according to documented procedures associated with the testing method.	
c	Uses procedures to screen test results to recognise improbable or unreasonable scores. Checks score scale-conversions and other clerical procedures for accuracy.	
d	Clearly and accurately labels scales and provides clear identification of norms, scales types, and equations used.	

The Learning Specification	
Essential Knowledge	**Essential Skills**
Test scoring procedures	Use of scoring keys and self-scoring test forms
Norms and standardisation	Use of norm tables
Aberrant response patterns	Conversion of test raw scores to standard scores
	Bringing together and documenting of tests scores and other assessment data for reports.
	Computation, where appropriate, of composite scores using standard formulae and equations.
	Management and filing of data
	Procedures for checking for clerical errors.

EWG4

European Test User Standards
for test use in Work and Organizational settings

Document 2: GLOSSARY OF TERMS[1]

A

ABILITY

ABILITY describes the degree to which someone can carry out certain types of mental operations -- generally operations which involve `reasoning' of some form.

ABILITY MEASURE *see* ABILITY TEST

ABILITY TEST

ABILITY TESTS vary in the types of operation they involve and the types of material they contain. Typical operations include 'analogies' (A is to B as C is to ?) and 'series completion' (1,2,4,8,?). The content of ability test ITEMS tends to concern words and sentences, numbers or shapes. Ability tests are generally designed to assess what people are capable of rather than what they have learnt or what they know.

ACCEPTABILITY

ACCEPTABILITY concerns those factors which make a test acceptable to the test taker and the test user. The notion of ACCEPTABILITY is important in the design and administration of assessment procedures and should be taken into account to ensure that the test taker cooperates in the procedure. ACCEPTABILITY is affected by the FACE VALIDITY of the test and by the test user's faith in it.

ACCURACY

The precision with which some attribute is measured. ACCURACY is a function of RELIABILITY or freedom from measurement error.

ACHIEVEMENT TEST *see* ATTAINMENT TEST

ADAPTIVE TESTING

see ADAPTIVE TEST.

[1] These definitions have been taken either from Bartram & Lindley (1994, 200), *Psychological Testing: The BPS Level A Open Learning Programme* or from Bartram & Lindley (2005). *Psychological Testing: The BPS Test Administration (Occupational) Open Learning Programme*, both published by BPS Blackwell. The contents of this Glossary are

ADAPTIVE TEST
A TEST where the items are selected from a large database or bank of items held on the computer. Each person who takes the test may be given a different selection of items, as the computer picks just those items that provide most information about that particular person's level of ability.

ADMINISTRATOR
The person who administers a psychological TEST.

ADVERSE IMPACT
The effect of a selection process which results in relatively more people from one group being selected than from another. ADVERSE IMPACT is one, but only one, of the conditions required to establish indirect discrimination. Also known as disparate or disproportionate impact.

APPARATUS TEST
A TEST that requires the manipulation of various items of specialised apparatus -- peg-boards, typing tests, etc.

APPLICANT
A person who has made an application for a position.

APTITUDE BATTERY
A sequence of APTITUDE TESTS which can be used to provide both detailed and general overall measures.

APTITUDE TEST
An APTITUDE is a potential to succeed at something in particular. ABILITY is assumed to underlie aptitude. APTITUDE TESTS are those which have been designed to measure those mental operations (or abilities) which affect the likelihood of someone acquiring some particular skill (for example, computer programming or TV repairing). ABILITY TESTS differ from aptitude tests in that the former are designed to assess the general reasoning skills while the latter tend to contain ITEMS with content which more specifically related to the aptitude concerned. However, the difference is largely one of function or use. In many cases, the same actual test may be used either as an ability test (to measure a person's general intellectual functioning) or an aptitude test (to assess their potential for success in some occupation). In practice, measures of GENERAL ABILITY can usually be drawn from aptitude test scores (especially APTITUDE BATTERIES).

ASSESSMENT
The process of appraising or estimating some attribute or set of attributes of a person.

ATTAINMENT TEST

The focus of attainment or achievement TESTS is knowledge and proficiency; on what has been learnt rather than on the ability to learn. These tests specifically assess what people have learnt and the skills they have acquired, for example, shorthand and typing tests.

AUTHENTICATION

One of the functions of test administration is to authenticate the identity of test takers. That is, to ensure that the person who presents themselves for testing is actually who they say they are. For high stakes testing, the test administrator should ask for some form of identification to confirm this. This should be some form of photo identification, not something that might have been given to an accomplice.

B

BIAS

BIAS occurs in TESTS whenever people's responses vary in some systematic way which is related to some characteristic which the test was not intended to measure. Factors which can produce bias in scores (at either the item level or overall test score level) include differences in sex, age, culture, educational background and literacy.

BREADTH

Assessment methods vary in both their BREADTH and their SPECIFICITY. A TEST of GENERAL ABILITY that samples several ABILITY DOMAINS may be regarded as `broad' and `general'.

C

CANDIDATE

A person who has taken part in the SELECTION process for a particular job or training course.

CLASSICAL TEST THEORY

A theoretical approach used in PSYCHOMETRICS that regards all observed scores as fallible and defines the relationship between observed scores and the TRUE SCOREs which are assumed to underlie them.

CLIENT

The person or persons who request a service (in this case, testing) either for themselves or as a

representative of an organization.

COGNITION
The internal processes and operations involved in perception, memory, thinking, reasoning and problem solving.

COMPETENCIES
Work-related behaviours that are instrumental in the delivery of desired result or work outcomes.

COMPETENCY PROFILE
A profile or list of the COMPETENCIES required for a job or position.

COMPOSITE TEST SCORE
A score produced by adding together scores of two or more tests or SUB-TESTS. In some cases, these may be differentitally weighted before they are added. COMPOSITE TEST SCORES are frequently produced by simply summing the RAW SCORES for each part of a TEST. COMPOSITE scores derived from BATTERIES of tests are often produced to provide a general measure of suitability in a selection situation.

CONFIDENTIALITY
Any information obtained using a psychological test should be considered as belonging to the test candidate in the first instance. Whatever is done with that information should be done only with their INFORMED CONSENT.

CONTROLLED MODE
This is a mode of test administration in which control is exercised over who can access a test on the internet and how often they can access it. It may also include controls over the location they can access it from and the time or date it is available.

CRITERION REFERENCING
In CRITERION REFERENCING a person's score on a TEST is used to predict or anticipate how they will perform on types of task not directly sampled by the test but which have been show to be correlated with test performance.

D

DATA PROTECTION and DATA PRIVACY LAWS
Legislation that gives individuals rights concerning personal data that is stored on computer or in other filing systems.

DECILE
A percentile-based scoring system where the RAW SCORES are divided into ten categories each containing 10% of the distribution.

DEXTERITY TEST
DEXTERITY TESTS are designed to assess various aspects of motor coordination, such as speed of movement, precision of fine motor control etc.

DISPOSITION
The temperament, personality, or characteristic mode of operating of a person.

DISTRIBUTION
In psychometric testing terms, the apportionment of scores obtained by people across all possible values of a variable. *See* FREQUENCY DISTRIBUTION.

DOMAIN
A universe, sphere or province of objects which meet some criteria. In relation to psychological testing people talk about ability TEST ITEMS being drawn, for example, from the DOMAIN of Verbal Reasoning, or from the domain of Spatial Ability. The term is also applied to areas of achievement. For example, mathematical attainment is a domain containing a range of skills involving arithmetic, algebra and so on.

DOMAIN-REFERENCED MEASURE
This is where CONTENT-RELATED VALIDITY data is used as the basis for interpreting a test score. The logic is to relate performance on a TEST to the level of performance required in a job by using a common quality standard. Judgements are then made about what level of performance on the test would be required for adequate performance of the job. The weakness of domain-referencing is that it relies on expert judgement. As such it really provides a way of generating hypotheses about how test performance should relate to job performance: it does not actually prove that the two are related. Where possible it needs to be backed up by CRITERION-RELATED VALIDITY studies.

DOMAIN REFERENCING
The process of relating a person's score on a TEST to levels of competence within some DOMAIN of knowledge or performance.

DRIVE
An attribute or need of a person which is considered to cause them to act in a certain way or motivate them to action.

E

ETHICAL (ISSUES)
Issues concered with the rights, responsibilities and obligations of those involved in testing -- the test taker, the test user, and the test user's client.

F

FACE VALIDITY
What, to the test taker, the TEST appears to measure. The superficial appropriateness of a test (see ACCEPTABILITY).

FAIRNESS
Fairness in testing is a relative term. A TEST is fair or unfair depending on whom it is used and how it is used. Its use is fair if it is not BIASED with respect to the groups with which it is used and if it can be shown to be valid. Thus, using either an unbiased test of mechanical reasoning or a biased clerical aptitude test in a clerical selection situation would be unfair. In both cases, differences between people would not be relevant for the job and hence it would be unfair to select people on the basis of such differences. Where a test is known to be valid but shows between group bias and the degree and type of test BIAS is known with respect to some selection procedure, then it is possible to practise fair selection by using different cut-off scores for each group.

FAIR SELECTION
The application of the principles of FAIRNESS to the selection process.

FEEDBACK *see* REPORTING BACK

FIVE POINT GRADING SCHEME
A common percentile-based scoring system where the top 10% of scores are classed as grade A; the

next 20% as grade B; the next 40% as grade C; the next 20% as grade D and the lowest 10% as grade E.

FREQUENCY DISTRIBUTION
The number of people who obtained each of the various values which could be obtained on a particular VARIABLE. A FREQUENCY DISTRIBUTION shows how people's scores are distributed across all possible values. FREQUENCY DISTRIBUTIONS are often used to examine the number of people obtaining each of the possible RAW SCORES on a TEST.

G

GENERAL ABILITY
ABILITY TESTS vary from those designed to give an overall measure of general intellectual functioning (GENERAL ABILITY TESTS) through those designed to assess broad areas of ability (for example, Verbal, Numerical or Spatial) to those focusing on specific MENTAL OPERATIONS (for example, three-dimensional spatial rotations). The latter tend to be used for aptitude assessment. General ability tests, in order to properly cover the full range of mental operations, tend to include ITEMS or SUBTESTS dealing with each of the main areas of ability. When GENERAL ABILITY is tested using a battery of ability TESTS, SPECIFIC ABILITY scores as well as an overall general ability measure can be obtained.

GENERAL ABILITY TEST
A TEST designed to give an overall measure of general intellectual functioning. *See* GENERAL ABILITY.

GENERAL INTELLIGENCE
The ability to perform on TESTS and in tasks which involve the understanding of relationships. The capacity to meet new situations, or to learn to do so by new adaptive responses.

GENERAL NORMS
General norms are intended to be representative of a large and diverse POPULATION. For example, UK general population; Norwegian adult males; French 16 to 18 year old school leavers. *See* NORMS and SPECIFIC NORMS.

GENERAL POPULATION NORMS
NORMS suitable for use with most people for converting RAW SCORES to either PERCENTILE or STANDARD SCORE based measures. Such norms are usually based on a large representative sample of people (in terms of age, sex and other VARIABLES).

GENERALISABILITY THEORY
This theory is concerned with how much we can generalize from scores obtained under one set of conditions to those we would be likely to obtain under another. This approach attempts to measure the effects of separate sources of ERROR VARIANCE which can affect the RELIABILITY of a TEST and then estimate the amount of ERROR that will be present in measures obtained under various different sets of conditions.

GENERATING REPORTS
The process of creating a report on a person's test results by using computer software. Computer-generated reports may be obtained either form stand alone PC applications or from service providers on the Internet

GRADES *See* FIVE POINT GRADING SCHEME

GUIDANCE
The process of giving information to a person to help initiate, support and clarify their decisions.

H

HISTOGRAM
A FREQUENCY DISTRIBUTION represented in a graphical form. Each score is represented by a bar, the height of which is equal to the number of people who obtained that score.

I

IDENTIFICATION OF DEVELOPMENT NEED
The use of tests to identify those areas of behaviour where a person may have potential for change. See PERSONAL DEVELOPMENT.

INFORMED CONSENT
The gaining of agreement from a test taker to the terms and conditions under which testing is to take place. This agreement must be obtained having first clearly informed the test taker of their rights and responsibilities, the reasons for testing, the type of tests to be used and what will be done with the results of the tests.

INSTRUMENT
A psychological TEST or other procedure for measuring differences between people.

INTEREST INVENTORY
An INTEREST INVENTORY is designed to assess, in a systematic manner, people's likes and dislikes for different types of work or leisure activity.

INTERESTS
Attitudes towards various types of activity, either in relation to work (vocational interests) or outside work.

IPSATIVE TEST
An IPSATIVE TEST compares a person's score on one scale with their own score on other scale(s). As a result, the scores on each scale are dependent on each other to some degree. Sometimes referred to as a self-referenced TEST.

ITC
The International Test Commission. This is a body that sets Guidelines relating to tests and test use through international consultation. Further details are on the ITC website www.intestcom.org

ITEM RESPONSE THEORY
A theory defining the relationship between test items and the likelihood of people making correct responses to those items in terms of the level of the `latent trait' (ability) they possess. The theory forms the basis for many ADAPTIVE TESTING systems.

ITEM SCORE
The numerical score given to a test taker's answer to an individual TEST item. For example, a correct response to an item in a numerical reasoning test may be scored as 1 and an incorrect response may be scored as 0.

ITEM
An ITEM is the smallest element within a test to which a score is assigned. Generally a question within a TEST, or a statement in a personality questionnaire.

J

JOB ANALYSIS
The process of formally defining the nature of a job, often in terms of the characteristics required by people who are to perform it.

JOB DESCRIPTION
A description of a job produced from the results of a JOB ANALYSIS

JOB-RELATED KNOWLEDGE ATTAINMENT TEST
A TEST designed to measure what a person knows in relation to a particular job.

JOB SIMULATION
JOB SIMULATION exercises are often used in the procedures that come under the general heading of the Assessment Centre Method. Job simulations may take the form of in-tray exercises, group problem-solving exercises and so on. They start from the assumption that the candidate does not yet possess the requisite knowledge or skill, but that the underlying ability will manifest itself when he or she works through an exercise that simulates the broad demands of the job in question. Job simulations are generally designed to assess APTITUDE rather than ATTAINMENT, though they may rely on the acquisition of some general job competences.

L

LEVEL OF DIFFICULTY
The level of difficulty concerns the degree of ABILITY required to answer a test ITEM. TESTS may either be designed to contain items of a similar difficulty level, or the level of difficulty may be increased as the test taker progresses through the test. The idea behind this approach is that it should provide a wider range of discrimination between people with the more able people getting further into the test.

LOCAL NORM
A particular type of specific NORM GROUP. This is a sample of people local to an organization. *Also see* NORMS.

M

MANAGED MODE
A mode of administration in which there is both direct supervision and control over the equipment being used, and other conditions. Typically managed mode administration refers to the use of dedicated testing centres.

MAXIMUM PERFORMANCE
Measures of MAXIMUM PERFORMANCE measure how well people can do things, how much

they know and how great their potential is. Measures of maximum performance include TESTS of ABILITY, APTITUDE and ATTAINMENT. These measures are usually distinguished from measures of TYPICAL PERFORMANCE which assess personality, vocational or occupational interests, needs, drives and levels of motivation.

MEASUREMENT
The procedure used to obtain a score from a person on some SCALE.

MEASUREMENT ERROR
Inaccuracies arising from the measurement process, the sources of which are due to extraneous factors which affect scores on a TEST. *See* SYSTEMATIC BIAS and RANDOM ERROR.

MENTAL OPERATIONS
MENTAL OPERATIONS are the internal processes and manipulations that have to be carried out to answer a TEST ITEM.

MODES OF TEST ADMINISTRATION
Classification by the International Test Commission of test administration modes into four types: OPEN, CONTROLLED, SUPERVISED and MANAGED.

MOOD INVENTORY
An instrument designed to assess a person's mood.

MOTIVATION
Factors which affect a persons likelihood of action and the choices they make between alternative courses of action. The reasons given as to why people perform certain acts. Motivation is often defined in terms of the goals which people seek to attain through their actions.

MULTIPLE-CHOICE (ITEM FORMAT)
For MULTIPLE-CHOICE items, test takers have to select one of a number of possible answers. Also see OPEN-ENDED (ITEM FORMAT).

N

NEED
A requirement which can act as the driving force or motivation for action. Needs include nurturance, affiliation, social approval, sex, self-actualization etc.

NORM GROUP
The sample of people from whom NORMS are derived. Also referred to as a REFERENCE GROUP.

NORM TABLES *See* NORMS

NORMAL CURVE *See* NORMAL DISTRIBUTION

NORMAL DISTRIBUTION
A symmetrical bell-shaped distribution with certain specific properties: the MEAN, MODE and MEDIAN are all equal to each other; the proportion of the values falling between any interval along the scale is known from the mathematical properties of the distribution. There will always be, for example, 68% of the values between -1 and +1 STANDARD DEVIATIONS. This form of distribution is found for a wide variety of both physical and psychological traits. Also called the `normal frequency distribution' or the `normal curve'.

NORMATIVE
Normative information included in a TEST's DOCUMENTATION enables the test user to see how a person's performance on the test compares with that of others.

NORMATIVE SCORE *See* NORM-REFERENCED MEASURE

NORM-REFERENCED MEASURE
A NORM-REFERENCED MEASURE defines where a person's raw score lies in relation to the scores obtained by other people (that is, a NORM GROUP).

NORM-REFERENCED SCORE
The score obtained on a NORM-REFERENCED MEASURE. Such scores are expressed either as some form of PERCENTILE SCORE or STANDARD SCORE.

NORMS
Information usually in the form of a table, which enables RAW SCORES to be converted into PERCENTILE SCORES or STANDARD SCORES (or both). Also see NORMATIVE, GENERAL NORMS and SPECIFIC NORMS.

O

OPEN-ENDED (ITEM FORMAT)

For OPEN-ENDED ITEMS, test takers have to write down their own responses to the items: Alternatives are not given. *See also* MULTIPLE CHOICE (ITEM FORMAT)

OPEN MODE
Open Mode is where the test taker has direct access to the test materials. So there is no involvement of a test user or test administrator. Such tests include the books of tests you might buy in the local book shop or the tests you can find on the Internet that are directly accessible to everyone. Often the only requirement is that you pay some money before you can access the test. However, no qualifications are required from you either in terms of test use or test administration.

ORAL FEEDBACK *See* REPORTING BACK

P

PERCENTILE *See* PERCENTILE RANK and PERCENTILE SCORE

PERCENTILE RANK
The value on the RAW SCORE scale below which a given percentage of the sample's scores lie. For example, if the 85th percentile rank is 16, then 85% of the sample will have scored less than 16. The PERCENTILE RANK is more commonly referred to as just the PERCENTILE or in some cases the CENTILE.

PERCENTILE SCORE
A number between 0 and 100 expressing a test taker's RAW SCORE in terms of the percentage of the norm group who scored less. *Also see* PERCENTILE RANK.

PERFORMANCE MANAGEMENT
Procedures used in organisations to help in managing the performance of employees towards the achievement of organisational goals and objectives.

PERSONAL DEVELOPMENT
Procedures by which individuals can develop their knowledge skills and COMPETENCIES. Testing may be used to help identify where their strengths lie and where they may have development needs.

PERSONALITY INVENTORY
Psychological TESTS that assess DISPOSITION, that is, preferred or typical ways of acting or thinking. Personality inventories attempt to measure how much or how little a person possesses of a specified TRAIT or set of traits.

PERSON SPECIFICATION
A description derived from a JOB DESCRIPTION of the personal characteristics necessary to do that job.

PHYSICAL MAKE-UP
For example, health, physique, appearance, bearing and speech.

POPULATION
A POPULATION contains all the people who conform to some specification. In PSYCHOMETRICS, normative reference groups are populations: for example, UK adult females; university arts graduates; general population. Psychometrics involves making inferences about people who come from some population on the basis of information known about the behaviour of a representative sample from that population.

POTENTIAL
A capacity or capability to perform or acquire the skills to perform some class of actions.

POWER TEST
The focus of a POWER TEST is on how many items a person is able to answer correctly. The time limit is designed to allow most people to complete all of the test ITEMS. If a person's score is mainly affected by their ability to answer the questions correctly -- rather than their speed -- the test is a power test.

PRACTICALITY
The notion of PRACTICALITY in TEST use concerns issues of cost-efficiency, such as what the test costs; what training is required to use it; how long it takes to administer, score and interpret; what equipment is needed.

PRIVACY
Ensuring that people's test data and other personal information is not disclosed to others without the INFORMED CONSENT of the test taker.

PROFESSIONAL (ISSUES)
Issues associated with professional practice and codes of professional conduct in relation to test use -- see also ETHICAL ISSUES.

PSYCHOLOGICAL TEST *See* TEST

PSYCHOLOGICAL TESTING

The use of psychological TESTS in the process of assessment.

PSYCHOMETRICS
Literally, the measurement of mental processes. PSYCHOMETRICS is the technology that underlies TESTS and their development.

Q

QUARTILE
A percentile-based scoring system where the RAW SCORES are divided up into four categories each containing 25% of the distribution (referred to as the first, second, third and fourth QUARTILES).

R

RAW SCORE
The RAW SCORE is the total number of correct answers a person obtains on completing a TEST. A raw score is the ABSOLUTE score a person gets on a test.

RECRUITMENT
The process of obtaining new hires into an organisation. In some cases, the term RECRUITMENT is used to refer to the whole process from applicant attraction through to hire. Others may use it to refer to the processes leading up to the definition of a short list from which final choices are made (SELECTION).

REFERENCE GROUP *See* NORM GROUP

REFERENCING SCORES
To compare a person's RAW SCORES on a SCALE against some other measure. The comparison may be with other people's scores on the same scale, the person's own scores on other scales, known relationships with other performance measures, expected levels of attainment in the domains from which the test itesm were drawn. *See* NORM-REFERENCED, SELF-REFERENCED, CRITERION-REFERENCED and DOMAIN-REFERENCED.

RELATIONSHIPS
TEST ITEMS where a test taker has to identify the relationship between two or more things and then use that relationship to select from a set of alternatives.

RELATIVE SCORE
A SCALE SCORE which describe a person's performance relative to that of other people or in terms of some other measure.

RELEVANCE
Another word used to describe the concept of VALIDITY.

RELIABILITY
The extent to which one can rely on the obtained TEST score being an accurate measure of a person's TRUE SCORE, rather than a measure of incidental random factors. RELIABILITY is usually assessed either by measures of INTERNAL CONSISTENCY, EQUIVALENCE or STABILITY.

REPORTING BACK
The process of feeding back the INTERPRETATION of TEST scores to a CLIENT or CANDIDATE. This can be in the form of oral feedback or a written report.

S

SAMPLING
The selection of a limited number of people (or other objects) from a defined POPULATION.

SCALE
In testing it is common to talk of measuring characteristics along a scale. Ability, for example, being a scale which goes from low to high scores. Thus scores obtained on a TEST of some characteristic are generally referred to as SCALE SCORES. *See* SCALE SCORES, RAW SCALE SCORE.

SCALE SCORE
The numerical scores attributed to a test taker's answers to individual TEST items are added up to provide a single measure called a RAW SCALE SCORE. Scores obtained on a test are generally referred to as SCALE SCORES.

SCOPE
In relation to psychological TESTS, the range of attributes covered by a test (the test DOMAIN) and the range of people with whom the test can be used.

SCORING
The process of marking the answers to a psychological TEST, including conversion of RAW SCORES into STANDARD SCORES.

SCORING KEYS
Tools for assisting in the marking of tests. For example, many scoring keys are in the form of templates that can be laid over the answer sheets.

SECURITY (OF TEST MATERIALS)
Guarding against the unfair and illegal use of TEST MATERIALS.

SELECTION
The process of choosing people with the best chances of succeeding in a job or on a training course.

SELF-REFERENCED MEASURES
SELF-REFERENCED MEASURES involve comparing a person's scores on one scale with their scores on other scales. *See* IPSATIVE TEST.

SELF-REPORT
SELF-REPORT measures are instruments that ask the respondent to answer a structured set of questions about themselves. Also called self-description instruments. Most personality and interest TESTS are self-report.

SPECIAL APTITUDE
Another way of describing a SPECIFIC APTITUDE, for example, mechanical, dexterity and so on.

SPECIFIC ABILITY
A particular ABILITY (for example, spatial ability).

SPECIFIC APTITUDE
A particular aptitude (for example, clerical accuracy).

SPECIFICITY
The degree to which a test assesses specific as opposed to general ability or aptitude.

SPECIFIC NORMS
NORMS which are based on some specific sample. For example, public sector engineering workers; clerical staff from a number of different companies. Also see NORMS and GENERAL NORMS.

SPEED TEST

A TEST which contains relatively easy ITEMS but which have a strict time limit. The measure of performance stresses the number of items attempted within the fixed time. *See* POWER TEST.

SPEEDED

It is common to see references to TESTS as more or less speeded. The more the standard time limit results in people failing to attempt some items, the more the test is said to be 'speeded'. *See* SPEED TEST and POWER TEST.

STANDARD SCORE

In PSYCHOMETRICS, the 'standard' scale developed for measuring psychological characteristics is called the Z-SCORE SCALE or sometimes simply the STANDARD SCORE SCALE. A z-score is a measure equal to one STANDARD DEVIATION (SD) of a distribution.

STANDARD SCORE SCALE

The most commonly used standard score scales are z-scores (one SD), T-scores (one tenth of an SD), STENS (half and SD), STANINES (half an SD) and IQ (generally one fifteenth or one twentieth of an SD depending on the scale used). *See also* STANDARD SCORES.

STANDARD ERROR OF MEASUREMENT (SE_m)

The amount of ERROR associated with making inferences about a person's TRUE SCORE from their obtained score.

STANDARDISATION

The procedure of establishing the initial set of NORMS for a TEST, defining the conditions under which it should be used, and of assessing its RELIABILITY and VALIDITY.

STANDARDISED

In PSYCHOMETRICS, a standardised measure, TEST or testing procedure is one which has known characteristics. For example, a CORRELATION COEFFICIENT is a standardised score (like a Z-SCORE) which is known to come from a distribution of possible values which range between zero and plus or minus one.

STATES

STATES are concerned with how a person is feeling or performing at a particular moment in time (for example, a current mood), rather than how they generally feel or typically perform. States are often distinguished from TRAITS which are more stable and enduring psychological characteristics.

SUPERVISED MODE

This is the mode in which the test administrator has direct face-to-face involvement with the test

taker. The test takers will come to a location where the test administrator is able to supervise them taking the test.

SYSTEMATIC BIAS
A source of MEASUREMENT ERROR which is predictable and can lead to possible unfair BIAS in the use of TESTS. SYSTEMATIC BIAS is potentially measurable.

T

TECHNICAL MANUAL The part of a TEST MANUAL that covers all the technical details of the test, such as design and development, RELIABILITY, VALIDITY, BIAS and STANDARDISATION

TEST
An assessment procedure designed to provide objective measures of one or more psychological characteristics. These include ABILITIES, APTITUDES, ATTAINMENTS, INTERESTS, beliefs, personality and so on. The important feature of psychological TESTS is that they produce measures obtained under standardised assessment conditions which have known RELIABILITY and VALIDITY. They provide a way of comparing a person's performance against that of others. An instrument that has been developed using psychometric principles. The term, test, is used as shorthand for psychological test or psychometric test, which includes various inventories and questionnaires.

TEST ADMINISTRATION
The process of administering a psychological TEST to one or more people.

TEST BATTERY
A sequence of TESTS. *Also see* APTITUDE BATTERY.

TEST DATA
The information about candidates and their scores resulting from the taking of a psychological TEST.

TEST INTERPRETATION
The process of attributing a meaning to a test taker's score on a psychological TEST, by reference to information about the test's validity, reliability and its NORMS.

TEST LENGTH

The number of ITEMS a TEST contains.

TEST MANUAL

The technical documentation accompanying a psychological TEST which tells the test user how to use and administer the test and what conclusions can be drawn from the results. Variously referred to as TECHNICAL MANUAL, USER MANUAL or TEST DOCUMENTATION.

TEST MATERIALS

The materials needed to administer and interpret a psychological TEST. These generally include a TEST manual, answer sheets, question booklets, profile sheets, administration instructions and so on.

TEST RELIABILITY

The RELIABILITY of a TEST.

TEST SOPHISTICATION

A level of awareness and knowledge of TESTS or testing without which a person's scores may be negatively biased. TEST SOPHISTICATION may arise from prior exposure, the process of testing, or through the use of practice tests and information describing testing procedures. This should be distinguished from active coaching in how to do a particular test which will unfairly inflate a person's scores.

TEST SESSION

The period during which a psychological TEST is administered.

TRAINABILITY TEST

A TRAINABILITY TEST is designed to see whether a person is likely to be able to cope with the training required to do a job. Typically, it consists of a highly-structured short training course with a test of performance at the end.

TRAIT

TRAITS are those relatively stable and enduring characteristics of people that make them predictable. Traits are usually distinguished from STATES which are a more changeable form of psychological characteristic, for example, current moods and feelings.

TYPICAL PERFORMANCE

Measures of TYPICAL PERFORMANCE are designed to assess disposition, such as personality, beliefs, values and interests and to measure motivation or `drive'. Measures of typical performance are usually distinguished from measures of MAXIMUM PERFORMANCE which are designed to assess how well people can do things and measure ability, aptitudes or attainment.

U

UNBIASED
Unbiased tests are those that do not show any systematic difference with regard to particular groups of people. BIAS is always relative to the groups concerned. A test might be unbiased with respect to age, while biased with respect to gender, for example.

USER MANUAL *See* TEST MANUAL

UTILITY
UTILITY concerns the benefits which accrue from the use of a psychological TEST. Utility is a function of the balance between issues of PRACTICALITY (the 'costs' associated with using a test) and RELEVANCE and FAIRNESS (the 'benefits' associated with the test). The benefits, in turn, are limited by the RELIABILITY of the test and its SCOPE.

V

VALIDATION
The process of building up evidence about what can and cannot be inferred from TEST SCORES.

VALIDATION OF RESULTS
In the context of test administration, this refers to the need to ensure that the results have been obtained under proper conditions and have not been the outcome of cheating and collusion.

VALIDATION STUDIES
Research studies that provide evidence on the VALIDITY of a test.

VALIDITY
Information on the validity of a TEST tells the user what inferences can be drawn about the person who has produced the score on a test and what is being measured by a test -- that is, what is it a test of. *See* CONTENT, CONSTRUCT and CRITERION-RELATED VALIDITY.

VARIABLE
VARIABLES are the characteristics which we attempt to measure with psychological TESTS. They are so named because they vary from person to person, or for the same person from time to

time. The defining characteristic of a variable is that a given person can have only one value of it at any one time -- for example, they cannot be 5' 5" AND 5' 10" at the same time. There are three main classes of variable: NOMINAL, ORDINAL, and SCALAR.

W

WORK SAMPLE

A WORK SAMPLE TEST is one in which the task has been taken from a job. The task is done under STANDARDISED assessment conditions. Work samples are essentially ATTAINMENT TESTS. They pre-suppose that the test taker has acquired some measure of a particular skill and sets out to see how much.

WORKPLACE COMPETENCIES

See COMPETENCIES

Z

Z-SCORE

A STANDARD SCORE scale with a MEAN of zero and a STANDARD DEVIATION of one.

EWG5

Document 3: European Test User Standards

For test use in work and organizational settings

Assessment Guidance for the UK Level A and Level B Qualifications in Occupational Testing

[DRAFT]

Standards prepared by

the European Federation of Psychologists' Associations
and
the European Association of Work and Organizational Psychologists

Assessment Guidance prepared by

The British Psychological Society

Version 1.8. 25 March 2005

Unit 1	Take responsibility for ethical test use
Standard: 1.1	Act in a professional and ethical manner
Standard: 1.2	Ensure you have the competence to use tests
Standard: 1.3	Take responsibility for their use of tests
Standard: 1.4	Ensure that test materials are kept securely
Standard: 1.5	Ensure that test results are treated confidentially

The Assessment Specification
Evidence required and rules of evidence
Levels A and B: Evidence demonstrating knowledge of general theories relating to testing and use of tests . Evidence demonstrating knowledge of chosen domain-related theories and knowledge Level B: A log book documenting involvement in all aspects of a psychological testing cycle. Certification or qualification evidence in relation to instruments utilised, as appropriate.

Standard: 1.1	Act in a professional and ethical manner	
The Assessment Specification		
Evidence that must be provided **The candidate should be able to:**		**Rules of Evidence**
Level A	Describe the main principles relating to the ITC Guidelines on Test Use and relevant local and national codes of Good Practice in Test Use	The assessor should require some positive evidence of compliance with the standard. It is not sufficient to argue that that the candidate has met the standard through lack of evidence to the contrary.
Level B	Present a Log book documenting evidence of having met the standards.	Log books could include, records of test usage, examples of testing policies developed with clients; emails or letters relating to good practice issues; evidence of checks carried out on suitability of employees etc. The assessor should require some positive evidence of compliance with the standard. It is not sufficient to argue that that the candidate has met the standard through lack of evidence to the contrary.

UNIT 1 Take Responsibility for ethical Test Use: Assessment Guidance

Standard: 1.2		**Ensure you have the competence to use tests**	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**Rules of Evidence**
Level A:			
A1.5	1.9	Describe the range of tests used in work and organizational psychology.	Examples should cover the range of types of test used in work and organizational assessment settings.
A1.3	1.7	Distinguish between norm-referenced, and other measures (e.g. mastery tests, workplace competence assessment procedures).	Show understanding of the difference between norm-referencing and referencing to some external criterion or standard.
A1.4	1.8	Describe what a work-sample test is and how it is used.	Include trainability testing and work simulations in this area.
A1.1	1.2	Describe the distinction between tests assessing 'maximum' performance and those assessing 'typical' performance.	This relates to the basic distinction between tests with right-wrong answers (ability, achievement, attainment measures) and those which ask people to rate or select options - self- and other report measures like personality, attitude, interests, 360 degree, climate measures and organizational surveys)
A1.2	1.3	Distinguish between tests of potential and tests of attainment.	Tests which are intended to predict what someone will be able to learn or do in the future and those which measure their current level of knowledge understanding or skill.
A	1.1	Describe the major theories of intelligence, differences between them and issues relating to them.	
A1.6	1.4	Describe the distinction between measures of general and specific ability.	Give illustrative examples of tests on the continuum from specific to general ability
A1.7	1.5	Describe how measures of ability are more or less influenced by environmental factors.	Genetic vs immediate environment. Long-term vs short-term stability
A	1.6	Describe how race, ethnicity, culture, gender, age, and disability may interact with ability.	
A3.4	4.1	Describe the basic premises of classical test theory	That actual measures are 'fallible' scores which contain a 'true' score and a random error.
A	4.2	Describe the basic premises of item response theory and its application in the field of work-related testing.	Not at the formal level of equations, but in terms of basic concepts
A	4.3	Describe what is meant by a test information function.	To understand that the amount of information obtained from a test varies with the ability level of the test taker, and that therefore, tests need to be tailored to ability levels if one is to maximize the gain in information for a wide range of test takers.
A	4.4	Describe the principle of adaptive testing	That test items can be selected on the basis of an estimate of an individual's ability in an interactive way, so that the test is optimally tailored to each person who takes it.
A2.1	3.1	Describe the concept of score distribution, measures of central tendency (mean, median, mode) and spread (range, SD).	Demonstrate understanding through ability to interpret histograms, bar charts etc
A2.3	3.2	Relate the mean and SD to positions on the measurement scale underlying a distribution of scores	
A	3.3	Describe the relationship between the mean, median and mode of a distribution.	
A2.4	3.4	Describe how the relative locations of mean, median and mode vary with the shape of the distribution.	
A2.5	3.5	Describe the ways in which the means and SD of samples may vary when they are drawn from the same population.	Guidance needed to explain 'population'
A2.6	3.6	Describe the relationship between the Standard Error of the mean of a sample of observations and the size of the sample.	To understand that the size of the error of estimation decreases as a function of the square root of the sample size.
A3.1	5.1	Describe what is meant by correlation.	Demonstrate understanding by being able to define the conditions under which the correlation coefficient is maximized (both positively and negatively) and is minimized and be able to interpret bivariate scattergrams

A3.5	5.2	Describe in outline the methods of estimating reliability and describe their relative strengths and weaknesses	Internal consistency (alpha), test-retest – same or alternate form, short or long time interval
A3.6	5.3	Describe why test scores may be unreliable	Measurement error, scoring error, situational factors, item sampling, etc.
Level B:			
B1.1	1.1	Describe the differences between implicit (personal, or everyday) and explicit (empirical, psychological) models of personality.	
B1.2	1.2	Describe the differences between temperament (personality), motivation (needs, interests and aspirations) values, attitudes, beliefs and ability.	
B1.3	1.3	Discuss ways in which genetic and environmental factors (such as culture and education) affect personality.	Broad overview of issues including concepts of stability in traits over time and development of personality through interactions of genotype and environmental factors
B	1.4	Discuss how race, ethnicity, culture, age and gender may interact with personality	
B1.4	1.5	Describe the key differences between the psychometric, psychoanalytic, social-learning, humanistic and behaviour analytic approaches to describing and explaining differences in personality.	
B1.5	1.6	Describe and illustrate the type and trait based approaches to personality and explain the limitations of each.	The distinction is between models that regard people as differing categorically from each other and those which regard there as being some underlying continuum. In some instances /traits/ are interpreted more narrowly as indicative of biological characteristics, in others they simply indicate within-person consistencies of behaviour.
B2.6	1.8	Describe the 'Big Five' factor model of personality and its importance	The Big Five provides an organizing framework for the domain of personality, as the concept of 'g' does for ability. Candidates should be able to give examples of how scales from two tests whic h were not developed as Big 5 measures can be mapped onto the five factor model
B2.1	2.1	Describe the advantages and limitations of the main approaches to assessing personality and give an example of the type of measure obtained from each.	
B2.2	2.2	Outline the main sources of and opportunities for 'sabotage' or 'distortion' in assessment, associated with each approach.	
B2.5	2.5	Describe the difference between trait and state measures.	The key difference being that traits are relatively stable and predictable, whereas states are more transient reflections of the impact of the current situation on people's behaviour. As examples of state measures one could include state anger/anxiety; measures of mood etc
B	1.7	Differentiate between trait-based descriptions and behavioural descriptions (competencies for example)	Trait-based descriptions make assumptions about the present of underlying or latent characteristics that explain the consistencies in behaviour, whereas competencies do not make this assumption.
B2.7	2.6	Describe ipsative scaling and ipsative item formats.	Be able to identify at least one instrument that uses this approach.
B2.9	2.8	Describe assessment procedures based on pseudo-scientific bases.	Give examples of pseudo-scientific assessment procedures which have high acceptability ('face validity'), are used in occupational assessment, or appear convincing to the lay-person but which have little or no proven validity. Illustrate the dangers associated with their use.

B6.1	5.1	Provide examples of instruments related to a range of different approaches to personality assessment.	Give examples of instrument based on the following approaches and describe how the instrument and approach are related: psychometric; psychoanalytic; social-learning, humanistic and behaviour analytic approaches to personality. Would expect projective techniques to be covered here [TAT, Rorschach, etc] at level of general knowledge not user expertise.
B6.2	5.2	Describe the position taken by situationalist theories.	Explain the implications of situationalist theories on the use and construction of personality measures and on the argument that generalized 'stable' dispositions can account for behaviour in different settings.
B6.3	5.3	Describe the role of personality theories in the design and interpretation of occupational interest inventories.	
B6.4	5.4	Identify self-report questionnaires based on each of the following approaches: type, trait, ideographic. Describe the advantages and disadvantages of each one.	Trait is used broadly to indicate instruments that assume continuous measures of characteristics that are normally distributed in the population. Ideographic approaches do not make assumptions about the general nature of the characteristics that are assessed (e.g. people have unique construct systems in Kelly's model)
B6.5	5.5	Describe how deliberate 'sabotage' and intentional 'distortion' can be controlled for.	Give examples of at least three different methods used to assess or control for these effects and describe how these are used in practice.
B6.8	5.9	Describe the difference between instruments that assume independence of traits and those that assume traits are correlated.	Give an example of one instrument which is based on traits which are assumed to be correlated with each other and one which assumes the underlying traits are independent.
B6.9	5.10	Describe the evidence supporting the 'Big Five' factor model of personality and give examples of scales from different tests which either do or do not fit the model.	
B6.10	5.11	Describe how profiles produced by ipsative instruments are affected by the number of scales on which they are based and illustrate with examples of two different tests.	This should show an understanding of how the degrees of freedom for scales to vary in an ipsative instrument are k-1 (where k= number of scales), while for normative instruments they are = k.
B6.11	5.12	Discuss in non-technical terms the main issues involved in the 'ipsative-normative' debate concerning when ipsative or normative measures should and should not be used.	This should cover the problem of factoring ipsative measures, practical advantages of ipsative measures (e.g. control over response bias and faking good) and issue of using norms with ipsative measures. Should also address the use of ipsative vs normative measures in selection and for use in development applications.
B6.6	5.6	Describe the general principles underlying factor analysis and its strengths and limitations as the basis for test construction.	In conceptual rather than mathematical terms
B6.7	5.7	Describe how instruments may be constructed using factor analysis or other rational techniques.	Give an example of one instrument which uses scales constructed using factor analysis and one which uses scales constructed using an alternative method (e.g. on the basis of item content).
B	5.8	Describe how type measures are constructed.	
B7.1	6.1	Describe the difference between dependability, stability and internal consistency.	Explain the difference between short-term and long-term retest reliability. That is long-term retest correlations may be low because the traits are not stable or because the measure is not dependable over time. Discuss the need for periodic reassessment of personality and why information about it may become 'outdated'.
B7.2	6.2	Describe how internal consistency is affected by the breadth of a characteristic and explain the relative advantages and disadvantages of highly homogeneous versus broad scales.	
B7.4	6.3	Describe the factors which contribute to unreliability in personality assessment	Distinguish between factors that affect variability in scores and factors which affect variability in the interpretations given to those scores.
B6	5.13	Demonstrate an awareness of current issues and debates in the area of personality assessment.	For example, debates over such constructs as emotional intelligence.

B8	7.2	Describe some of the novel ways in which computer-based assessment can be carried out	Use of multimedia, realistic task-based exercises; IRT scaling of graded response items etc.
B9	8.3	Discuss how internet testing has affected assessment practice	With reference to at least three of the following assessment functions: (a) Selection; (b) Promotion; (c) Redundancy/outplacement; (d) individual personal development; (e) team development; (f) career guidance; (g) counselling

UNIT 1 Take Responsibility for ethical Test Use: Assessment Guidance

Standard: 1.3		Take responsibility for their use of tests	
The Assessment Specification			
Evidence that must be provided **The candidate must be able to:**			**Rules of Evidence**
Level A:			
A7.2	12.4	Ensure that all mandatory requirements relating to the rights of candidates and clients and obligations under relevant current legislation are clearly explained to both parties.	Legislation includes the Data Protection Act 1998 for the UK.
A6		Provide feedback of information about results to the test candidate which:	
A6.12	11.3	Provide the candidate with opportunities to ask questions, clarify points and comment upon the test and the administration procedure and comment on the perceived accuracy and fairness or otherwise of the information obtained from the test;	
A6.14	11.4	Clearly inform the candidate about how the information will be presented (orally or in writing) and to whom.	
A	11.9	Deal sensitively with scores lying outside the candidate's expectation and provide necessary support and guidance.	
Level B:			
B5	4.7	Take responsibility for the final report, whether written by the test user or computer generated	

Standard: 1.4		Ensure that test materials are kept securely	
The Assessment Specification			
Evidence that must be provided **The candidate must be able to:**			**Rules of Evidence**
Level A:			
A7.4	12.3	Ensure that all test materials are kept in a secure place which is not accessible to people other than authorized test users.	
A7.5	12.5	Ensure that potential test candidates are not provided with prior access to test materials other than those specifically designed to help candidates prepare for their assessment.	

Standard: 1.5		Ensure that test results are treated confidentially	
The Assessment Specification			
Evidence that must be provided **The candidate must be able to:**			**Rules of Evidence**
Level A:			
A7.1	12.1	Ensure that best practice is followed in giving clear descriptions to the candidate(s) prior to their assessment concerning: how their results are to be used; who will be given access to them; for how long they will be retained.	
A7.3	12.2	Ensure that all test data are stored in accordance with current legislation, in a secure place and access is not given to unauthorized personnel.	Ensure understanding of relevant aspects of Privacy, Digital Rights Management, EOC Legislation in the country of use and test taker's rights. Test User must know how and where data is stored, if in a digital environment. If Test User delegates some part of the process to someone else, they must be competent to exercise responsibility.

UNIT 2 Follow good practice in the use of tests: Assessment Guidance

Unit 2	Follow good practice in the use of tests
Standard: 2.1	Evaluate the potential utility of testing in an assessment situation
Standard: 2.2	Choose tests appropriate for the situation
Standard: 2.3	Give due consideration to issues of fairness in testing
Standard: 2.4	Analyse and interpret results appropriately
Standard: 2.5	Communicate the results clearly and accurately to relevant others
Standard: 2.6	Review the appropriateness of the test and its use

The Assessment Specification
Evidence required and rules of evidence
Knowledge of the general theories and models. Knowledge of chosen domain related theories and knowledge. Demonstration of practitioner competence in real or simulated scenarios. Log book documenting involvement in all aspects of a psychological testing cycle. Certification or qualification in relation to specific instruments utilised.

Level B: [NB references are to pre 2005 Unit numbers]	
Foundation LB Unit 2. Personality assessment	
Range: Coverage of approaches to personality assessment should include at least 5 of the following: (a) self-report procedures; (b) reports by others, as in 360 degree feedback; (c) group or individual situational assessments; (d) projective measures; (e) observations of behaviour; (f) task performance measures (that is, 'objective' measures of personality); (g) physiological measures; (h) ideographic methods (e.g. repertory grids); (i) multi-task assessments; (j) social-psychological approaches.	
Test Use: LB Units 3 and 4 (Module B2 for first instrument and B3 for subsequent instrument(s))	
Competence relating to Test Use should be assessed in relation to specific, substantive instruments. The focus of the assessment is on the assessee's competence as a practitioner to demonstrate a good understanding of the strengths and weaknesses of the instrument, to make appropriate use of it in different assessment contexts and to weight information obtained from it appropriately with other information about the candidate.	
Test Use Units 3 and 4 [old Level B numbering] require evidence of competence in interpretation and providing feedback to both clients and candidates in each of two types of context: (a) Client-oriented, assessment contexts (where the prime reason for the assessment is to provide the client or client organization with information about the candidate): a. report to a client organization for selection, promotion or inclusion/exclusion in a team, redundancy/outplacement; b. feedback to candidate(s) on the above. (b) Candidate-oriented assessment contexts (where the prime reason for the assessment is to provide the candidate with information about him or herself e.g. Individual personal development, Team development, Career Guidance, Counselling): c. report to a client on the development/career potential etc of a candidate; d. report to candidate(s) on the above.	
It is realized that candidate-oriented situations may arise (a) where it is not appropriate to provide feedback to the client organization, or (b) where the distinction between client and candidate does not apply as they are the same person. The evidence obtained must cover both of these types of context and include instances of verbal feedback to a candidate and instances of written feedback both to a client and to a candidate. The evidence should be collected under realistic working conditions.	
Test Choice and Evaluation LB Units 5, 6, 7 and 8 [Old numbers] (Module B4)	
These Units cover the knowledge and understanding deemed necessary for test users to make more informed and in-depth choices between instruments as to their suitability for various purposes. The focus in Unit 6 [old number] on test construction assumes a level of knowledge sufficient to make an informed evaluation of an instrument on the basis of reviews and information presented in its technical manual. The level of understanding and the practical skills required to design and construct tests are not required for Level B.	
Unit 6 [old number] extends the coverage of reliability and validity issues dealt with in Level A and – as with all the Level B Units – assumes an underlying competence at Level A.	
For Unit 8 [old number], evidence should be provided that the assessee has a knowledge base on which to make choices which is broader than that covered by the specific instruments on which their Test Use Units are based.	

UNIT 2 Follow good practice in the use of tests: Assessment Guidance

Standard: 2.1		**Evaluate the potential utility of testing in an assessment situation**	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
A	2.1	Explain the purpose of job analysis and competency profiling.	
A	2.2	Describe the differences between structured and unstructured methods of performing job analysis and competency profiling.	Give examples of each type of approach (for example, job analysis questionnaires competency profiling tools, interviews with job incumbents, etc) and describe when they might be appropriate.
A	2.3	Describe the practical aspects of doing job analysis or competency requirements analysis in an organization, and how to deal with related practical challenges.	Include access to job subject matter experts, obtaining sufficient data etc.
A1.8	2.4	Describe the relationship between job analysis, competency profiling or other procedures for defining assessment criteria and the production of a person specification for a job or role.	Understand the difference between job descriptions and person specifications
A	2.5	Explain the rationale for deriving assessable person attributes from the person specification.	
A1.9	2.6	List, for each characteristic of a person specification, one or more possible methods of assessment.	How specific should methods be?
A1.10	2.7	Describe how knowledge of occupations and information on work performance are important for the proper use of tests in career guidance.	
A	7.1	Describe the relationship between base rate, selection ratio and validity.	Show how average quality of selected applicants is a joint function of all three of these factors.
A4.10	7.2	Describe how one would assess the utility associated with using psychological tests as part of some personnel decision-making process.	Issues around estimation of dollar criterion, validity of assumptions about use of top-down selection etc.
A4.11	7.3	Describe the way in which information about a test (especially predictive validity data) can be used to reduce the risks associated with personnel decisions.	Explain how validity measures convert to utility gains
A	7.4	Apply standard utility equations to make a business case for the relative values of different validity solutions.	Cronbach-Brogden-Gleser equation. General rules of thumb for estimating dollar criterion values etc.
A	7.5	Illustrate the economic utility of a systematic approach to job analysis and competency profiling prior to recruitment	Be able to provide examples in the form of cases studies that illustrate gains in utility (reduced turnover, increased profit, increased quality of work etc)
Level B:			
B9.2b	8.2.2	Describe what corroborative information about personality and/or interests could be collected using other methods.	With reference to at least three of the following assessment functions: (a) Selection; (b) Promotion; (c) Redundancy/outplacement; (d) individual personal development; (e) team development; (f) career guidance; (g) counselling
B9.1a	8.1.1	Discuss factors which limit the practicality and appropriateness of using various different types of instrument (ipsative, non-ipsative; self-report, behavioural etc).	Evaluate in relation to two instruments with reference to at least three of the following assessment functions: (a) Selection; (b) Promotion; (c) Redundancy/outplacement; (d) individual personal development; (e) team development; (f) career guidance; (g) counselling
B9.1b	8.1.2	Discuss how and why information about personality and/or interests should and should not be used.	Evaluate in relation to two instruments with reference to at least three of the following assessment functions: (a) Selection; (b) Promotion; (c) Redundancy/outplacement; (d) individual personal development; (e) team development; (f) career guidance; (g) counselling

B8.1	7.1	Discuss issues of test equivalence, reliability and test-taking attitude in relation to differences between paper-and-pencil and computer-based test administration procedures.	This should included discussion of the distinction between supervised and unsupervised computer-based and Internet modes of administration.

UNIT 2 Follow good practice in the use of tests: Assessment Guidance

Standard: 2.2		**Choose tests appropriate for the situation**	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
A	6.1	Describe the central importance of construct validity in establishing the validity of a test	Be able to describe how all other forms of validity provide aspects of construct validation.
A3.13	6.2	Describe and illustrate the distinctions between face, faith, content, construct, criterion-related and consequential validity.	
A3.16	6.3	Describe the relationship between reliability and validity.	Explain why it is impossible to have higher validity than reliability.
A	6.8	Describe in outline the procedures used in meta-analysis and describe the key findings from validity generalization research	Show an understanding of the basic concept of correcting for artifacts and aggregation of results across studies.
A3.17	6.9	Describe the conditions under which one may use evidence form meta-analyses to support the use of a test in some new situation.	Understanding of the results of validity generalization analyses and an understanding of what credible intervals mean. Also being aware of the conditions under which situational specificity can and cannot be ruled out.
A4.2	8.1	Use test publishers' catalogues, specimen sets and other reference materials to identify one or more instruments potentially suitable for a particular function.	This should cover a range of sources, including test reviews as well as catalogues.
A4.3	8.2	Identify, for each test, information in the test manual which relates to the test's construction, rationale, reliability, validity, its norms and any specific restrictions or limitations on its areas of use.	
A4.4	8.3	Identify relevant practical considerations.	Ease of administration, time required, special equipment needed, etc.
A4.6	8.4	Compare information presented about the test's validity with relevant aspects of the assessment specification and make an appropriate judgment about their fit.	Person specification; job description
A4.7	8.5	Make a suitable judgment about the appropriateness of norms, benchmarks or reference groups in terms of representativeness and sample size.	
A4.5	8.8	Examine any restrictions on areas of use and make an appropriate judgment as to whether the test could be used.	E.g. age; cultural or ethnic limitations; ability range, etc.
A4.12	8.11	Make a final selection of test(s) that is defensible in terms of the available evidence about the appropriateness of psychological testing given the assessment specification.	[also ITC 1.3]
Level B:			
B7.5	6.4	Provide a reasoned consideration of the evidence for validity associated with two substantially different instruments in relation to their construct validity.	This should cover all aspects of evidence relating to construct validity - including criterion-related evidence and content validity
B7.6	6.5	Discuss two substantive criterion-related validity studies associated with personality or interest assessment instruments used in occupational assessment.	
B9.2a	8.2.1	Identify and justify the use of a personality and/or interest assessment instrument.	With reference to at least three of the following assessment functions: (a) Selection; (b) Promotion; (c) Redundancy/outplacement; (d) individual personal development; (e) team development; (f) career guidance; (g) counselling

UNIT 2 Follow good practice in the use of tests: Assessment Guidance

Standard: 2.3		Give due consideration to issues of fairness in testing	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
A4.8	8.6	Identify whether use of the test would meet the mandatory requirements of the national and international legislation and identify whether use of test would break the law.	For the UK, this includes Equal Opportunities (Sex Discrimination), Race Relations, Disability Discrimination and Data Protection Acts, plus other EU Directives.
A4.1	8.7	Follow the law relating to direct and indirect discrimination on the grounds of gender, age, sexual orientation, community group or disability in recruitment and selection.	Both national laws and EU directives
A	8.9	Follow best practice in testing in relation to ensuring fairness of outcome for members of minority or potentially disadvantaged groups	Need to describe what is good practice in relation to these and how practice relates to the law and policy in the local country
A4.9	8.10	Describe best practice regarding when adjustments to the test are necessary before test administration to accommodate for specific disabilities.	Modify test administration instructions to accommodate disabilities. The importance of not compromising the test's technical qualities should be given due regard. With reference to technical recommendations and restrictions regarding the test (including copyright), the assessee should show they can decide on the specific adjustments, including a recommendation not to use, that could reasonably be made to a test's administration to accommodate any disability encountered.
Level B:			
B2.3	2.3	Describe how problems of bias in personality assessment can be controlled and/or limited.	Including language, cultural and ethnic group bias
B2.4	2.4	Explain how problems of bias also apply to informal methods of personality assessment (e.g. in everyday social interaction, in interviews and so on).	
B7.11	6.10	Discuss how current legislation and guidelines relating to fairness in the use of tests affect the use of personality measures in occupational settings.	This should include reference to group-based score differences, differential validity, impact on use of cut-scores and cultural or language issues.
B9.2c	8.2.3	Discuss how current legislation and guidelines relating to fairness in the use of tests relates to the use of the [specific] instrument.	With reference to at least three of the following assessment functions: (a) Selection; (b) Promotion; (c) Redundancy/outplacement; (d) individual personal development; (e) team development; (f) career guidance; (g) counselling

Standard: 2.4		**Analyse and interpret results appropriately**	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
A	6.4	Describe the basic principle of combining multiple measures to create a single predictor score.	Explain how uncorrelated measures will, together, provide a better overall prediction than correlated ones, given similar levels of correlation between each measure and the criterion. Be familiar with the concept of multiple regression and know what 'R' indicates.
A	3.7	Describe the differences between raw-scores, and standardized scores.	Give illustrative examples of each type of scale.
A	3.9	Describe the differences between point scores, banding and ranking of candidates.	
A6.2	10.1	Make an informed choice about whether and how to use information provided in the manual about norms or cut-off scores	Select appropriate norms tables, where available and attach suitable cautions to interpretation of the results; or not use the test, where no relevant norms or cut-off tables are available.
A	10.2	Describe the implications of using separate norms for people belonging to different groups (e.g. race or gender).	
A3.7	5.4	Describe how reliability is affected by changes in the length of a test.	Understanding of the Spearman-Brown prophecy formula.
A3.8	5.5	Describe how reliability is affected by range restriction.	Impact on values obtained in different studies, with more or less homogeneous samples of people; importance of looking at the SEm
A3.9	5.6	Describe how different levels of confidence are computed from raw and standard scores using the standard error of measurement.	Be able to explain why confidence limits increase as the level of confidence required increases, and how this is related to the Standard Error.
A3.10	5.7	Describe how the standard error for the difference between two scale scores or for the sum of two scale scores is related to the standard error of measurement of the two scores.	Be able to explain examples. Importance of understanding this when looking at the difference between people on the same scales or differences over time in a person's performance on the same scale.
A3.11	5.8	Describe how the degree of correlation between two scale scores affects the reliability of (a) their sum, and (b) the reliability of the difference between them.	
A6.7	10.6	Provide interpretations of scale scores paying due regard to the correlations which exist between each pair of scales.	Interpretation need to take account of the extent to which different scales are correlated partly because of the redundancy introduced by correlation and also because of the impact of correlation on the standard error of their difference.
A6.8	10.7	Make appropriate connections between performance on a test and characteristics in the original person specification.	Demonstrate the ability to relate test scores back to the assessment specification requirements in a way that will be intelligible to a lay person.
A	10.8	Describe the relationship between group score differences, selection ratio, validity and adverse impact.	Understand that lower cut=scores have less adverse impact, other things being equal, and that adverse impact is a function of how a test is used not of the test itself.
A	10.9	Take into account the impact of any accommodations for disability in interpretation	For example the impact on standard error of measurement.
Level B:			
B4.3	3.3	Discuss the issues involved in choosing suitable norm groups or reference groups for the interpretation of scale scores.	Can comment on the effects of using: norms based on broad based samples versus those based on narrow ones (small variance); mixed gender or ethnic group versus single gender or ethnic group norms; occupationally-related versus general population norms. Where type measures are concerned, the user should understand how to interpret patterns in relation to distributions of types in various reference groups.
B7.9	6.8	Describe the differences between actuarially and clinically derived composite scores based on weighted combinations of scales	For example, adjustment specification equations; criterion-referenced prediction equations; and so on. Need to show a clear appreciation of the difference between those which are actuarial (based on empirical data) and those which are clinical (conceptual or based on judgments about scale contents).
B2.8	2.7	Discuss the dangers of reliance on spurious face validity.	For example, 'Barnum effects' - using the candidate's self-assessment of the validity, literal interpretation and acceptance of scale labels.

UNIT 2 Follow good practice in the use of tests: Assessment Guidance

B4.1	3.1	Outline in non-technical terms the rationale underlying the instrument and the approach adopted.	
B4.2	3.2	Interpret tables of correlations with due regard to the numbers of scales and prior hypotheses	Not over-interpret 'chance' significance in tables of correlations produced from an instrument with a large numbers of scales.
B4.4	3.4	Interpret scale score profiles with due regard to the technical qualities of the instrument.	That is, reliability and validity. This also requires proper use of any built in 'validity' checks, lie scales, social desirability measures etc.
B4.6	3.5	Relate the interpretation given to a score appropriately to the information provided about the instrument being used and to the nature of the scale and its mode of construction.	
B4.7	3.6	Corroborate information about personality obtained from the instrument with information obtained from other sources.	
B4.8	3.7	Make appropriate use of any other relevant information (such as reports from other assessments of the candidate, self-assessments, peer-assessments etc) to aid the interpretation of results.	
B5.1	4.1.1	Demonstrate sufficient knowledge of the instrument to provide competent interpretation	To produce balanced written reports for at least two assessments in each case for: (a) the candidate, and (b) the client – where the assessment is being carried out for a third party.
B5.4	4.4	Produce written reports which provide a contextualized and overall balanced appraisal of the information available about the person.	Reports integrate the information on personality with that on ability and other relevant aspects of the person and present this within the context for which the information is sought.
B5.5	4.5	Ensure that neither organizational nor personal decisions are based solely upon the interpretation of data from a personality inventory.	
B7.7	6.6	Discuss the relative merits of information obtained about the validity of a scale from validation against local criteria and from validity generalization meta analysis.	
B7.8	6.7	Describe issues associated with the interpretation of (a) linear and (b) non-linear relationships between individual personality scales and external criteria.	
B7.10	6.9	Discuss why it is important to corroborate information about personality obtained using one type of instrument with information obtained from other sources to avoid misinterpreting 'method-variance' as 'trait-variance'.	It is important in order to avoid misinterpreting method variance as trait variance.
B8.2	7.3	Compare the relative advantages and disadvantages of computer-generated interpretative reports with those written by human experts.	In relation to reliability, validity, acceptability, cost and usability.
B8.3	7.4	Describe in outline the main techniques used by computer-based report generators to produce their reports.	
B8.4	7.5	Describe how the validity of computer-generated personality reports can be formally assessed.	This should include studies of the validity of the conclusions made in the report; the ability of reports to differentiate between people in a reliable and valid way. Reference can be made to the BPS/EFPA report review criteria.
B	4.6	Provide a critique of computer generated reports to identify where modifications might be needed to take account of feedback and to improve contextualization.	

B8.5	7.6	Outline the practical and professional issues associated with the use of computer-generated reports:	Reports should be reviewed for suitability and accuracy by the test user before being passed on to any third party. Issues need to be discussed in relation to reports as (a) an aid for the test user; (b) a report to the client; and (c) a report to the test take (candidate).

UNIT 2 Follow good practice in the use of tests: Assessment Guidance

Standard: 2.5		Communicate the results clearly and accurately to relevant others	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
A6.4	10.3	Represent the candidate's scores appropriately in terms of its reliability and comparability the scores of others.	Give due consideration to the comparability between the candidate and any reference groups, the standard error of the group mean and the standard error of measurement of the candidate's scores.
A6.5	10.4	Present norm-based scores within a context which clearly describes the range of abilities or other relevant characteristics of the norm group they relate to.	[compare old LA 11.2]
A6.6	10.5	Describe the meanings of scale scores in terms which are consistent with the construct validity evidence, which reflect the confidence limits associated with those scores and which are intelligible to the client and the candidate.	
A	11.1.0	Provide feedback of information about results to the candidate which:	
A6.10	11.1	Is in a form appropriate to his or her understanding of the tests and the scales;	
A6.11	11.2	Describes the meanings of scale names in lay terms which are accurate and meaningful;	
A6	11.5.0	Provide written reports for the client and/or candidate which:	
A6.15	11.5	Present in lay terms the rationale and justification for the use of the test.	
A6.16	11.6	Describe the meanings of scale names in lay terms which are accurate and meaningful;	
A6.17	11.7	Explain any use of normed scores in relation to the ability range of the norm group	
A	11.8	Justify any predictions made about future performance in relation to validity information about the test;	
A6.18	11.10	Give clear guidance as to the appropriate weight to be placed on the findings where appropriate.	
Level B:			
B5.1	4.1.2	Demonstrate sufficient knowledge of the instrument to provide competent oral feedback	To at least two candidates in each case
B5.2	4.2	Discuss results with candidates in a non-judgmental way.	Methodical use of the feedback interview to help confirm/disconfirm hypotheses generated from the pattern of individual test results.
B5.3	4.3	Provide an indication to the candidate and to the client (when there is a third party involved) of the status and value of the information obtained and how it relates to other information about the candidate's personality.	

Standard: 2.6		**Review the appropriateness of the test and its use**	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
A3.14	6.5	Describe the procedures used to assess concurrent and predictive criterion-related validities and explain the pros and cons of each procedure.	
A	6.6	Describe the problems of obtaining good criterion measures and the implications of this for criterion-related validity studies.	
A3.15	6.7	Describe how measures of validity can be affected by selection effects and range restriction.	
A	7.6	Describe how systematic validation of test use may generate value for the user organization	Be able to explain the relationship between validity and utility and give examples of how gains in validity might be translated into gains in utility.

Unit 3	Follow good practice in the administration of tests
Standard: 3.1	Make necessary preparations for the testing session
Standard: 3.2	Administer the tests properly
Standard: 3.3	Score the test results accurately

The Assessment Specification
Evidence required and rules of evidence
Provide a Log book documenting involvement in administration of tests in a variety of settings within a chosen domain.

Standard: 3.1		Make necessary preparations for the testing session	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
A7.1	12.1	Ensure that best practice is followed in giving clear descriptions to the candidate(s) prior to their assessment concerning: how their results are to be used; who will be given access to them; for how long they will be retained.	
A5	9.1	Distinguish between open, controlled, supervised and managed modes of test administration	As defined by the ITC guidelines on CBT and the Internet, Open mode is free access on the web; controlled mode requires the test taker to receive a unique login and password (often referred to as remote administration), supervised requires the presence of a test administrator and managed involves the use of a secure test centre.
A5	9.2	Describe the main issues associated with administering different types of tests in each mode.	Candidates should be able to describe the potential problems of candidate authentication and collusion in open and controlled modes of assessment. They should also be able to differentiate between issues of control associated with ability tests (timing, item security etc) and self -report inventories.
A5	9.3	*For controlled mode assessment*:	I.e. where the test taker is not under direct supervision, but they have been provided with a unique (generally one-time) login.
A5	9.3.1	Ensure that the tests being used are suitable for use in this mode.	Has the publisher provided evidence to support use of the test in this mode, or developed it specifically for this mode of administration?
A5	9.3.2	Ensure that candidates are sent the necessary instructions and log in details in good time.	The instructions should include a clear explanation about the purpose of the assessment, how the data will be treated, privacy and confidentiality, and provision for how to let the test user know if internet access is a problem for the test taker.
A6	9.3.3	Ensure that candidates understand what is required of them and what will be happen after completion of the assessment and that they have a means of raising queries.	For example, the importance of establishing an honesty contract with candidates in a selection procedure and explaining to them that scores obtain in unsupervised situations may be checked later in the process.
A	9.3.4	Describe the potential problems of collusion in remote assessment	
A	9.4	*For supervised and managed mode assessments:*	Those modes where a test administrator is present during the test session.
A5.1	9.4.1	Ensure that the location is suitable for the duration and type of assessment, equipment required and numbers of candidates	Arrange seating and desk space to minimize the possibilities of cheating. Inform the candidates of the time and place well in advance and ensure they are adequately prepared – where relevant – for what they will be required to do and why.
A5.3	9.4.2	Ensure, where re-usable materials are being used, that they are carefully checked for marks or notes which may have been made by previous candidates.	
A5.2	9.4.3	Ensure that any equipment (e.g. computers) is operating correctly and sufficient test materials are available for use by the candidate.	
A	9.4.4	Ensure that any necessary test accommodations for disabled candidates are available, well understood and appropriately provided.	

Standard: 3.2		**Administer the tests properly**	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
		For supervised and managed mode assessments:	Those modes where a test administrator is present during the test session.
A5.5	9.4.5	Brief candidates on the purpose of the test session and key rules and put them at their ease while maintaining an appropriate atmosphere.	Rules will include not talking to or colluding with others during the session, not having any recording devices, including mobile phones - which should be turned off and left with the test administrator.
A5.11	9.4.6	Deal appropriately with any questions which arise without compromising the purpose of the test.	
A5.12	9.4.7	Follow standard test procedures and instructions as specified in the user manual.	Provide the candidates with sufficient time to work through example test items. Make careful checks to ensure proper use of the answer sheet and response procedures. Explain any time limits and ensure that during the test candidates maintain silence to avoid distracting others. Make clear that once the test has begun no further questions can be answered. Has adhered strictly to the directions and instructions as specified in the test manual while making reasonable accommodations for persons with disabilities. Adhere strictly to test-specific instructions concerning pacing and timing.
A	9.4.8	Monitor behaviour during the test to identify any issues or problems that may arise.	Issues might include candidates who are attempting to cheat or record information about the test; candidates who are confused or not following instructions etc. Has ensured potential sources of distraction (e.g., wristwatch alarms, mobile phones, pagers) are removed; that test-takers have the materials they require for taking the test before it begins. Has ensured that test takers are not left unattended or subjected to distracting activities during a supervised test session.
A5.6	9.4.9	Ensure that all necessary information has been logged and recorded including details of any provisions that were made for any special needs.	Ensure that the candidate's personal details have been recorded, together with relevant details of what assessment instruments were used, what accommodations were made for any special needs and whether accommodations met their needs. Record any other relevant information – including notes on any particular problems which arose during the session which might have affected a candidate's performance.
A5.15	9.4.10	Check and collect all relevant materials at the end of the test session, and ensure they are stored securely.	Collect all materials when each test is completed check that all materials have been recovered and lock test materials away in a secure place. For P&P tests this includes test questions and answer sheets. For all tests this should include any notes or records made by the candidate during the test session. Visually check paper answer sheets for ambiguous markings which could be obscured by scoring keys or cause problems with machine scoring systems.

Standard: 3.3		**Score test results accurately**	
The Assessment Specification			
Evidence that must be provided **The candidate should be able to:**			**'Rules' of Evidence**
Level A:			
A5.21	9.4.11	Demonstrate accurate use of a range of different hand-scoring keys and 'self-scoring' forms for paper and pencil testing.	
A2.11	3.8	Use norm tables to look up percentile equivalents of raw s cores and various standard scores.	Standardized scores should include those passed on percentiles and those based on z -score conversions (T-scores, Stens etc)
A	9.4.12	Make appropriate records of candidate's raw scores and score conversions or check the accuracy of these records where this is process is automated.	
7.3	12.2	Ensure that all test data are stored in accordance with current legislation, in a secure place and access is not given to unauthorized personnel.	Ensure understanding of relevant aspects of Privacy, Digital Rights Management, EOC Legislation in the country of use and testee's rights. Test User must know how and where data is stored, if in a digital environment. If Test User delegates some part of the process to someone else, they must be competent to exercise responsibility.

EWG6

Consultation Questionnaire: Aggregated responses as of 1 May 2005

Please complete as much as you can of the following questionnaire once you have reviewed the relevant documents.

Please give or email the completed questionnaire to your EFPA-EAWOP Working Group representative or email your completed response to Prof Dave Bartram, the convenor of the EFPA-EAWOP Working Group: Dave.Bartram@shlgroup.com

Your details

Q1

Respondents:	Four from Sweden One from Norway Two from Finland One from Germany Two from UK

Q2: I am responding:

	On behalf of the following European organization, commission or association: [please specify]
5	On behalf of the following National organization, commission or association: [please specify] *STP* *STP* *Norwegian Psychological Association* *Finnish Psychological Association and The Finnish Work and Organizational Psychologists Association* *BPS Steering Committee on Test Standards UK*
3	On behalf of the following company, organization or group: [please specify,] *The Finish Psychological Corporation* *SHL Group*
2	On my own behalf.

Q3: If responding on behalf of a specific committee or group within an organisation, please specify:

I/We have the following involvement in testing [tick as many as apply].

6	Professional association of psychologists
1	Professional association of non-psychologists who make use of tests
4	Trainer or training organisation for test users
3	Test user or organisation of test users
	Test taker, group or organisation representing test takers
4	Developer of tests
4	Publisher/distributor of tests
3	Other [please specify] *Sub editor at the Swedish organization "Stiftelsen för tillämpad psykologi" (STP). PhD Student of psychology at Swedish Lund university.* *Test User Quality Body* *Academic*

Q4: Would you like your contact details to be added to our mailing list so that you can be notified of further developments in relation to this project?

10	YES
	NO

Consultation Questions

Structure of the standards:

	Very unclear	Unclear	Adequate	Clear	Very clear
Q5: How clear is the way in which the Standards have been structured?				7	3

If unclear, please suggest how might this be improved?	*The basic structure is good; it covers well all the phases of assessment work. Some minor sharpening is needed to avoid overlaps and repetitions*

Content of the standards

	Much too little detail	Too little detail	About right	Too much detail	Much too much detail
Q6: Is the amount of detail provided about right?			8	2	

Are there any specific areas where either more or less detail is needed?	*Some parts could be edited a little bit, e.g. St 2.3 or 3.1*

	Very poorly	Poorly	About right	Well	Very well
Q7: In general, how well does the draft Standards cover the areas of competence required of test users in work and organizational settings?			5	3	2

Q8: Are there any specific aspects of competence that are missing?

2	YES
8	NO

If YES, can you define what they are?	*In Finland we do not make such a distinction between psychometric and non-psychometric methods. From our viewpoint the standards should cover all methods used in assessment work*

Q9: Are any aspects of the competences that are specified not necessary?

2	YES
8	NO

If YES, can you say which ones?	*I'm slightly unsure about the statement in section 1.1:e that it is a required performance to represent tests and testing in a positive manner in communication with the media. Is it really a sound requirement to require of test users that they need be positive in contacts with the media? Couldn't realistic be enough of a requirement in that respect?* *Knowledge about Item-Response-Theory: Basics should be sufficient.*

Qualifications related to the standards

	Not all valuable	Not valuable	In between	Valuable	Very valuable
Q10: How valuable do you think test users would perceive a European accredited qualification in test use to be?			2	6	2

Q11: Do you think that test user qualifications in your country should be consistent with this European Standard?

10	YES
	NO
If NO, can you can you say why not?	

General comments related to the standards

Q12: If you have any specific comments to make on the content of the draft Standards, please provide those here.	*An overall approach to the problem is missing, for instance human psychology vs. behaviour psychology* *Test users responsibility* *National and international research* *More cooperation between test publishers* *Standard 3.1 e: this requires communication skills* *Due to my experience as a lecturer for Psychological Assessment and Selection, I find it extraordinarily important for students and young professionals to thoroughly "rehearse" test feedback. Communication of results - especially via orally given feedback - is an utmost difficult thing to do. Training should involve "hot" situations, due to the fact that not every result is desirable. Euphemisms do not help here.* *Another issue is related to the aforementioned: socio-economic backgrounds of test use should also be taken into consideration. We do not test aliens in outer space, but people within social environments which have a more or less direct impact on test feedback (over 5 Mio people in Germany are unemployed. Negative test outcomes within a selection process should be reflected on this background).* *One applauds the commitment and awareness that has driven the draft to this stage.* *Clarity and Useability* *- table entries/headings that are hard to unpack or potentially ambiguous because of the idiosyncratic use of initial capitals and punctuation. Eg: in lists where the end is not clear; or following a leading clause meant to act as stem to a number of subsequent sentences.* *- the use of Standard identifiers, eg 1.1, 1.2, in the describing the Unit make-up but which are not used when referring in later sections to the same Standards.* *- the ubiquitous appearance of "Test user standards" after Unit/Standard headings as if it were a further sub-heading, when it is the overall heading and so need not be repeated (or, if it is to be, not in that inferior position).*

	- *the repetitious way of presenting the material (which can be disorientating or can lead to small changes being missed). Nowhere is this more the case than in the descriptions of "Occupational context". One full description of the contexts would be vastly more helpful than repeated versions so abbreviated as to downgrade their intelligibility. Any variations from this could then be highlighted. In contrast entries under "Performance required" communicate much more effectively.* *Content and Coverage* *Terms such as "appropriately", "necessary", "properly", are used regularly without definition of the factors or the degree of these that must underlie the judgements that these terms imply. They could, for example, mean different things in different countries although apparently working to the same standards. I see the difficulty in pursuing definitions down to a truly behavioural level, that is perhaps inherent in the competence approach, but the alternative requires some de facto consensus on what is, say, appropriate if it can remain unanalysed. Could something be added to address this point?* *Phrases like "in an assessment situation" are used where the sense seems to be "in a particular assessment situation" rather than any situation. "Testing people" is contrasted on p9 with "testing groups" so it perhaps refers to "testing individual people".* *Instances of Essential Knowledge include "Classical test theory" and "Item response theory". This seems too sweeping. Even the substance of older books on these topics, such as by Gulliksen and Lord & Novick, would be known by relatively small proportions of test users. Maybe it could read "The approaches taken by .... ".* *Approach* *A drawback of the approach taken in the draft is that it appears to lose sight of what the assessment process is about, of what is being delivered by this means, eg effective psychological measurement that helps life decisions that are having to be made by or about people. It is the quality of such decision outcomes that is paramount; the Standards represent simply a route chosen to achieve this quality. Would there be any room for a comment about such core delivery aims? Or about indicators of poor delivery?*