

ATP SA RESPONSE:

South African Guidelines on Computerised Testing

Herewith the response of the Association of Test Publishers of South Africa (ATP SA) in reference to the Professional Board for Psychology's request for feedback on the South African Guidelines on Computerised Testing (dated 23 May 2006, Ref: 18/11/B; posted on the 'Open for Discussion' page of the Board's website). ATP SA appreciates the opportunity provided to give input on the Board's proposed guidelines for computerised testing. We would like to emphasise that we, as test publishers, undertake to act professionally, ethically and in a responsible manner.

Table of Contents

Background and History of the Association of Test Publishers	1
General Standpoint of ATP SA Concerning Computerised/Internet Testing in South Africa	3
ATP SA's Detailed Commentary on the <i>South African Guidelines on Computerised Testing</i>	5
Preamble	5
1 Clarification of Terms	11
2 Requirements Regarding Computerised and Internet-Delivered Tests for South Africa	11
2.1 Classification process and research evidence	11
2.2 Supervision and administration	13
2.3 Ethical considerations	13
2.4 Age restriction for respondents	13
2.5 Access control	14
2.6 Computer-generated reports	14
3 Give Due Regard to Technological Issues in Computer-Based (CBT) and Internet Testing	14
3.4 Consider reasonable adjustments to the technical features of the test for candidates with disabilities	14
4 Attend to Quality Issues in CBT and Internet Testing	15
4.1 Ensure knowledge, competence and appropriate use of CBT/Internet testing	15
4.2 Consider the psychometric qualities of the CBT/Internet test	16
4.3 Where the CBT/Internet test has been developed from a paper and pencil version, ensure that there is evidence of equivalence	17
4.4 Score and analyse CBT/Internet testing results accurately	18
4.5 Interpret results appropriately and provide appropriate feedback	18
4.6 Consider equality of access for all groups	20
5 Provide Appropriate Levels of Control over CBT and Internet Testing	21
5.1 Detail the level of control over the test conditions	21
5.2 Detail the appropriate control over the supervision of the testing	21
5.3 Give due consideration to controlling prior practice and item exposure	23
5.4 Give due consideration to control over test-takers authenticity and cheating	24

6	Make Appropriate Provision for Security and Safeguarding Privacy in CBT and Internet Testing	26
6.1	Take account of the security of test materials.....	26
	Appendix A: ATP SA Response	27
	Appendix B: Mode of Administration and the Stability of the OPQ32n: Comparing Internet (controlled) and Paper-and-pencil (supervised) Administration.....	28
	Appendix C: Mode of Administration and the Stability of the OPQ32i: Comparing Internet (controlled) and Paper-and-pencil (supervised) Administration.....	30
	Appendix D: Testing on the Internet: A validity study of ability tests using Item Response Theory (IRT).....	34
	Appendix E: Applicant Perceptions of the Fairness of Online Assessment in the Selection Process.....	36

Background and History of the Association of Test Publishers

The Psychometrics Committee met with developers and distributors of psychometric tests on 8 April 2005, at Johannesburg International Airport, following which the test publishers formed a group to place suggestions to the Board as part of the consultation process. This resulted in the formulation of the Association of Test Publishers of South Africa (ATP SA).

ATP SA is currently in discussion with ATP International. There is already a European division and they are now busy with the creation of an African division. ATP International was established in 1992 in the USA. It is a non-profit organisation representing providers of tests and assessment tools and/or services related to education, employment, certification/licensing or clinical issues. The mission of the industrial division is:

"Business organisations use a variety of tests as aids in hiring, placing or promoting employees. Industrial/Organisational test publishers are dedicated to promoting and advancing the role of quality assessment in the workplace."

The mission statement of ATP International was accepted for ATP SA and was sent to the Board, indicating that the Test Publishers Association of South Africa accepts these principles and is in the process of negotiating a South African affiliation with ATP.

While the Industrial Division of the association was already formed in 2005, the HPCSA requested that the association's membership be expanded to span all test publishers in South Africa. This was subsequently done, inviting all test publishers included on the distribution list of the HPCSA to a meeting held on 19 June 2006.

This document presents ATP SA's comments on the proposed *South African Guidelines on Computerised Testing*. The content was drafted based on individual members' commentary and sent out to all test publishers on the distribution list of the HPCSA for approval. The following test publishers have read this final proposal to the Board and subscribe to the recommendations in the document:

Mr Nic Laycock	Amos Laycock Consulting (Distributor)
Dr Terry Taylor	Aprolab
Ms Lisa Ashton	Bioss
Dr Louis Fick	Corporate Solutions Consulting
Ms Dawn Philip	Insights
Ms Frances McKelson	International Training Network
Dr Jopie de Beer	Jopie van Rooyen
Dr Renate Scherrer	Jopie van Rooyen
Dr Marie de Beer	M&M Initiatives
Ms Theresa Cotterrell	MAC Assessment & Development
Dr Maretha Prinsloo	Magellan Consulting
Ms Talita Kriek	Mindmuzik Media
Mr Frikkie Kriek	Mindmuzik Media
Dr Nicolaas Claassen	Mindmuzik Media
Ms Dalray de la Harpe	NMMU Student Counselling
Mr Hendrik Bronkhorst	Organisation & Management Technology
Dr Pieter Bronkhorst	Organisation & Management Technology
Mr Ben Venter	PISA
Ms Laura Kartus	Placements Inc.
Dr Pieter Erasmus	Potential Index Associates
Mr Colin Wilford	Profiles International
Mr Mark Cunningham	Profiles International
Prof Hennie Kriek	SHL
Ms Tina Joubert	SHL
Ms Kim Dowdeswell	SHL
Mr Chris Schutte	Thomas International
Prof Vasi van Deventer	Thomas International
Richard Sale	EDAC (Cyprus); Participant in the development of the ITC <i>International Guidelines on Computer-Based and Internet Delivered Testing</i>

Only Ms Nanette Tredoux of Psytech SA requested her name to be removed from the final document submitted to the Board.

The document has two sections, namely ATP SA's general view concerning computerised/Internet testing in South Africa, and a detailed commentary on the guidelines proposed by the Board.

General Standpoint of ATP SA Concerning Computerised/Internet Testing in South Africa

- [1] ATP SA agrees that computerised/Internet-based testing should be conducted in accordance with best practice guidelines. The ITC *Guidelines on Computer-based and Internet-based testing* provides a valuable reference for best practice that is the product of an intensive process of international collaboration and review. ATP SA recommends that the international best practice be utilised and fully agrees with the ITC guidelines. There are however differences of opinion as to the HPCSA interpretation and application of the guidelines. It is believed that the points contained in the international guidelines are valid for the South African context as well and that essential components of the international guideline have been removed in the SA guidelines, namely unsupervised testing. Without this component there is no need for these requirements.
- [2] A concern held by ATP SA is that the legal context in South Africa has been misinterpreted in the guidelines published by the Board. This is evident in the classification and control of psychological tests as well as the use of psychological tests, regardless of the platform used for delivery. Pertaining to this, ATP SA has requested Senior Counsel to formulate a legal opinion reflecting the correct interpretation of the relevant provisions. The legal opinion will be forwarded to the Board as soon as it has been received.
- [3] The proposed South African guidelines do not recognise the Open / Controlled modes of administration, limiting the use of CBT/Internet testing to the Supervised and Managed modes. One of the participants of the ITC drafting process, Richard Sale (EDAC), noted that this is "a potentially damaging twisting of the intent of the ITC". The question must be asked what the justification of this restraint is, as currently no explanation is given in the proposed guidelines?
- [4] ATP SA has previously made the point that establishing rules and regulations taking into account all complexities when testing across different platforms would be unfeasible and impractical (see ATP SA's *Response: Reclassification of Psychological tests on the current list of*

classified tests, submitted to the Board on 30 June 2006 and included here in Appendix A). Additionally, ATP SA recommended implementing the Test Review Model of the European Federation of Psychologists' Associations (EFPA), included in Appendix A, which covers the intended mode(s) of use or conditions under which the instrument was standardised and validated. This model also accommodates Internet and computer based testing, and therefore the adaptation of the ITC guidelines for SA can be questioned.

According to this approach, an instrument's appropriate modes of administration would be identified during the test classification process, regardless of whether the instrument was delivered via paper-and-pencil, computers and/or the Internet.

ATP feels that there is a need for a holistic approach to the practice of testing in general. The classification of tests, skills and competencies required to use the different categories of tests as well as the training requirements to use different categories of tests should all form part of this broader approach. The HPCSA adaptation of the ITC guidelines on Computerised Testing is not in line with the intent of the original document of the ITC, and all computer and Internet based tests should be classified as part of the broader classification system of Psychological tests. The use of computers and the Internet is just another medium in the collection and presentation of assessment data and should form part of the integrated system of test classification.

ATP SA's Detailed Commentary on the *South African Guidelines on Computerised Testing*

Herewith follows ATP SA's detailed commentary on the South African Guidelines on Computerised Testing. In the following sections the general views of ATP SA regarding computerised/Internet testing are supported and expanded on, with specific reference to the Board's proposed South African guidelines.

Preamble

- **Preamble Paragraph 2**

The proposed guidelines state "The Health Professions Act clearly reserves the use of psychometric tests for the profession of psychology."

ATP SA respectfully reminds the Board that allowance is made for other professional groups to use tests as specified in Act 56 of 1974. We also remind the Board that numerous stakeholders have requested this issue to be discussed as the reality is that many other professions require metrics to advance their own sciences and most/all professions in the Human Sciences have an interest in tests. It is also for this reason that the HPCSA has often been requested to define exactly what a Psychological test is and when a test is not seen as Psychological.

Test publishers commented that the most controversial part of the South African Guidelines on Computerised Testing (SAGCT) is the preamble, and specifically the argument that persons administering tests have to be "registered with the Board" (p. 3). It should also be noted that the term 'test' is used here as an abbreviation for the more general idea of psychological assessment.

The first point to be made is that the requirement of registration is out of line with international test practice. It is clear that the guidelines suggested by the Professional Board for Psychology ignores two modes of administration of tests and this is out of pace with what is done in other parts of the world.

This approach makes it difficult for South Africa to partake in the global economy. Tests are not only developed internationally, they are also used internationally (international companies operate in SA), and psychological testing plays an ever-increasing role in the workplace. South Africa cannot afford to be isolated from these practices. This does not mean that tests used in South Africa should not be adapted (there are international guidelines for test adaptation), but it does mean that South African organisations must be able to apply the international business models of test use. There are clear indications that the HPCSA's strict measures of control could obstruct the application of internationally accepted business models of test use, which can only serve to isolate South Africa from the rest of the world.

The fact that the fundamental psychometric properties of testing (e.g. reliability, validity and non-bias) are the primary guarantees against any testing malpractice (including actions of oppression) is the international basis of best practice in Assessment. The Employment Equity Act already includes these measures of protection for the public. Legislating these properties simply implements an academic and ethical principle. It simply says: "Let's agree that we will lawfully bind ourselves to stick to proper academic and ethical principles in testing". Therefore this legislation cannot be used as an excuse to tighten control over and above what internationally would be considered proper academic and ethical practice.

- (a) ATP SA fully endorses the Employment Equity Act and appreciates the role the HPCSA has played in the formulation thereof. It would be particularly useful if the HPCSA could refine the definition of a Psychological test so as to ensure that it is very clear when a test is psychological and when it is not. As Psychologists and Test Publishers we fully recognise the vulnerability of the disadvantaged, but also recognise that testing is used by many and assessment should be fair to all.
- (b) The argument that the South African population is in need of special protection that exceeds international practice actively maintains a paternalistic discourse that perpetuates the past. Unnecessary levels of control make it more difficult and also more expensive for members of the public to access psychological services. Does the HPCSA really serve the interest of the public when it is biased to rather withhold than to promote psychological services?

The resolution by the Professional Board that "... all tests must be administered by persons registered with the Board ..." and this includes "... highly structured and computerised tests ..." (p. 3) clearly demonstrates a bias towards the withholding rather than the promoting of psychological services. This is not in line with international practice; it would increase the costs of psychological testing and it may lead to organisations reverting to unreliable techniques and subjective practices (e.g. interviewing).

The reason provided for this resolution is that registration ensures levels of training, and ethical and professional conduct. No one would dispute the fact that professional registration is required to guarantee the level of professional practice, especially to oversee the broader implementation and use of assessments in industry. Different categories of tests, however, need different levels of qualifications of test users. International as well as local research clearly indicates that for example some personality tests can be administered remotely over the Internet, without a supervisor, without any loss of psychometric integrity (See Appendix B, C, and D).

What is disputed is the fact that each and every activity in the administration of a test requires knowledge, skills and ethical decision-making at the level of a registered person. One needs psychological knowledge to select an appropriate test, but one certainly does not need psychological knowledge when one follows prescribed instructions to prepare a test room, to welcome people to the room, to hand them a questionnaire to complete, to explain the purpose of the questionnaire, etc. Exactly how is the good of the public served by the requirement that a highly trained (and therefore expensive person) is to be used to perform a structured and easily encapsulated set of activities that does not require knowledge, skills and ethical decision making of a professional level?

- **Preamble Paragraph 5:**

"It is recognised that the duty of being in charge of a test room entails considerable responsibility, which goes beyond the mere ability to administer a given instrument. Not all respondents are equally able to deal with computerised testing, and the test administrator may sometimes have to

recognise this, and recommend an alternative mode of administration, if this is available.”

ATP SA acknowledges the concept of standard assessment procedures as core to the Psychology profession, we also remind the Professional Board that to register as a Psychologist at least 6 years of training has been done – as such some allowance for professional judgement and responsibility needs to be made. The same reasoning would be where assessment is done and the literacy level of a candidate is less than required for a certain assessment. Issues like these are real, but you cannot formulate rules or laws for every exception. This is why the Psychologist, ultimately carries the responsibility for his/her professional judgement.

- **Preamble Paragraph 6:**

“Important, life-changing decisions are made in test rooms (the guide reads restrooms!!), and the person in charge needs to be able to deal with many eventualities in a professional way.”

Most of the important, life-changing decisions associated with testing are not made in the test room. They are made before and after testing; before when tests are selected, and after when information gathered from a number of different sources is collated. Contexts that require profound decision making during testing are normally unstructured, and obviously require psychological knowledge, skills and ethical decision making actions

- **Preamble Paragraph 7:**

“With computerised testing and reporting systems, narrative reports are easily produced. This yields highly confidential information that needs to be dealt with professionally. Modern computerised systems are rarely used in such a way that the person administering the tests has no access to the results.”

Firstly, one does not have to be registered to be able to respect the confidentiality of information and to know how to treat confidential information. The ability to handle a document should not be confused with the ability to consume and apply the information contained in the document.

The knowledge, skills and ethics required to treat information confidentially are not psychological.

Most of the well designed and modern computer systems have very sophisticated security measures to protect results and test materials and can actually be better protected than paper and pencil materials. Computerised and internet testing are in most cases set up in such a manner that reports are made available only to designated recipients. The online systems or software systems of reputable publishers are typically set up in such a way that reports are made available to the professional person in a hard copy format or sent to the confidential email address of accredited, registered and trained persons.

- **Preamble Paragraph 8:**

“Trained observation of the behaviour of persons undergoing tests is sometimes a valuable source of information for the person reporting the results. This requires more skill than the mere administration of a structured test.”

Behaviour during testing can be a source of additional information, but this is important in unstructured and clinical settings. It is seldom the case in structured and group testing, especially in the work context. Note that the main difference between Individually administered tests and Group administered tests is that with group testing, monitoring of individual behaviour is NOT possible. This is a very basic principle of assessment. In most cases, computer-administered and internet-based assessment will be more similar to the typical group test administration situation and should not be compared to what is typically required and expected in an individual test administration situation. In the latter, the focus is often more diagnostic as opposed to the primarily assessment-related focus of group assessments.

While the benefit of group test administration is the cost- and time-effectiveness, it is generally understood that this does not allow for the monitoring of individual behaviour – which forms part of typical individual assessment.

In conclusion, the nature of the testing (being individual / unstructured versus group / structured) plays an important role in determining the potential value of additional information obtained from test behaviour observation. This again highlights the need for a more comprehensive and holistic test classification system.

- **Additional Comments on the Preamble:**

It is noted that previous documentation stated:

“Where the administration and scoring of computerised psychological tests is highly structured, a psychologist may decide to use a suitably trained assistant to administer and score the test under supervision.”

The current document has put in place a system with different – and seemingly higher levels of training and registration required for all aspects related to computerised assessment.

There seems to be some confusion and lack of clarity with regard to the practical implementation of the suggested guidelines – in particular when considering not only the individual assessment or more clinically-oriented environment. Some recommendations are suggested which may address some of the main concerns.

Taking into account the South African context, recognition of prior learning and SAQA guidelines was also raised by test publishers. In most business contexts, the recognition of prior learning has been recognised. The general reasoning behind this approach is that individuals who may not have been in a position to obtain formal qualifications, but who have solid experience in their particular field, should be given recognition for their expertise – which allows for them to perform certain functions for which they have proven experience and recognised levels of expertise. In contrast to this trend, the reverse has been the case in the field of psychology, with stricter requirements in terms of formal training being required and no recognition of practical experience given. Whereas previously, a general category of A-level test user was available – with limited training in Psychology required – this category as well as that of Psychotechnician, has been dropped and the new guidelines for registration are only for Psychologists and Psychometrists.

ATP SA draws the attention of the HPCSA to the fact that as governing body with the goal of guiding the profession of Psychology and protecting the public, more energy should be spent on managing and controlling the availability of software assessment packages or the thousands of internet assessments developed and distributed nationally and internationally, often of doubtful quality. As test publishing and distributing firms, managed and populated by registered psychologists, our commitment is to stay within the ethical guidelines of the profession. The purpose of our dialogue with the HPCSA is to draw the attention to the realities of international best practice and the fact that whilst Psychologists and test publishers are being prosecuted, those members of the public who are not registered are able to sell tests without any concern to their profession, the public or control.

1 Clarification of Terms

The wording of some of the terms – e.g. computerised report, computerised test and Internet-delivered test – appears to contain value judgements in addition to objective definitions. It is recommended that only objective definitions of terms be included in this list. In addition, the glossary provided at the end of the ITC’s guidelines is more in-depth, covering a wider range of terms that the reader may not be familiar with.

The proposed South African guidelines defines a ‘Test User’ to be “A person registered with the Health Professions Council of South Africa” to administer the test in question. This does not make allowances for the potential outcomes of ATP SA’s recommendation to re-evaluate the overall process of test classification in South Africa. The line manager or end user is also a test user.

2 Requirements Regarding Computerised and Internet-Delivered Tests for South Africa

2.1 Classification process and research evidence

We once again request the HPCSA to distinguish between a Psychological and a non-psychological test and accommodate this in test classification. In addition, for the evaluation of internet-delivered tests (as well as any other tests) clarity

needs to be obtained in terms of a guarantee from the HPCSA that the HPCSA nominated evaluators of these test:

- have the necessary qualification and experience to do such evaluations. It is not clear why the identity of test evaluators should not be made known to the test developers/publishers.
- should sign a document (together with the HPCSA) to ensure that confidentiality, copyright and ownership of the assessment will be (refer 6.1 this document) respected and adhered to at all times.
- committing to a reasonable time in which the evaluation and feedback will be finalised.
- will send back all the submitted copies of the assessment to the distributor/developer and as such at least limit some of the costs incurred in the HPCSA required registration process. The costs involved in the evaluation process are very high and often one of the reasons why developers/distributors cannot send all their tests in for evaluation to the HPCSA. Please keep in mind that the market of SA test publishers and distributors is limited to Psychologists and Psychometrists, that the publishers/developers sponsor all South African research on the products they develop or import (no government grants as in the previous dispensation), that 3 copies of all components of tests need to be submitted to the HPCSA as well as having to pay the HPCSA evaluation fee- with no guarantees from the HPCSA in terms of speed of evaluation, adherence to copyright, confidentiality or who maintains the material in his/her possession.

As has been requested in previous correspondence with the HPCSA, ATP SA strongly favours the EFPA test review format over the procedure currently in use. The EFPA's Test Review Model approach means that the classification covers the intended mode(s) of use of an instrument. The focus of the test classification process, therefore, is on the test itself and not on a delivery platform.

Where the proposed guidelines read "For tests that are found to be psychological tests a classification certificate will be issued specifying the categories of professionals who may administer the test and who may interpret and give feedback on the results", ATP SA recommends that definitions of psychological and non-psychological tests be included.

2.2 Supervision and administration

Globally, the trend in regulating the use of psychological instruments is to empower the psychologist to take control of the total psychological testing process. This allows the psychologist to use his / her professional judgement in determining what appropriate controls are, etc.

There are many different kinds of assessment available on internet or software. We do agree that there are a range of computer and internet tests that should be supervised but one cannot regard all tests as the same! As mentioned, research indicates that some tests can be administered without any supervision without any loss to the psychometric qualities of the test. New developments also use different hurdle assessment strategies to control or verify unsupervised internet assessment results. We do need to allow for professional judgement.

This section in the proposed guidelines finishes off by saying "Psychologists in charge of assessments for industry may not require prospective candidates to complete unsupervised psychological tests via the Internet." This statement is against all international developments and a first in the world!! The original version of the ITC was specifically designed to help guide professionals in this new domain of unsupervised assessments via the internet. By removing this option there is no need for internet guidelines.

One of the reasons behind the development of internet based assessments is to increase accessibility. Career guidance can now be done via the internet in remote areas. It is not always possible or cost effective to have a registered person present everywhere where testing is done.

2.3 Ethical considerations

ATP SA fully agrees with this section.

2.4 Age restriction for respondents

ATP SA would appreciate the rationale behind the statement saying that no Internet-Delivered psychological services may be rendered to clients under the

age of 18 years. Children, having grown up with computers are often much more comfortable and advanced in the use of computers and the internet than many adults.

2.5 Access control

ATP SA assumes an electronic signature or documented proof of the accountable Psychologist that signed off on the cut scores or algorithms used must be included in the software. In the industrial context cut scores are determined by the psychologist but implemented by the HR practitioner. In some software the cut scores are determined by the Psychologist but implemented in the software.

2.6 Computer-generated reports

This point concerning the use of computer-generated reports and their dissemination to assorted users does not take into account the context within which the report is to be used. While the proposed guidelines' points may be in line for clinical instruments they are not taking into account the industry context where reports are designed specifically for end users such as line managers. Such reports are generated using actuarial formulas that are compiled using the professional judgement of a psychologist, research, and applied en masse to the work context.

3 Give Due Regard to Technological Issues in Computer-Based (CBT) and Internet Testing

3.4 Consider reasonable adjustments to the technical features of the test for candidates with disabilities

In Point ii under Test Users, it is felt that the ITC guidelines' reference to the International Guidelines on Test Use adds value to the overall document, and should be retained. This point should then read:

"Follow best practice as in other modes of testing (see ITC Guidelines on Test Use). Ensure that any necessary test modifications specifically address

the test-taker's special needs and are within acceptable limits so as to not adversely affect score validity."

4 Attend to Quality Issues in CBT and Internet Testing

4.1 Ensure knowledge, competence and appropriate use of CBT/Internet testing

Test Publishers

[vii] The proposed South African guidelines remove reference to test publishers' documenting the limitations of the test in terms of the professional context in which it operates, which is a valuable point that should not be discarded. As such, ATP SA recommends this point be included as Point vii under the Test Publishers section:

"For Internet testing, document the limitations of the test in terms of the professional context in which it operates:

- *provide a statement indicating the limitations of the relationships between test user and test-taker that can be achieved through this mode (e.g. the Internet is a impersonal medium and a test user may provide only limited advice)*
- *provide a statement stating that there are limitations to the conclusions that can be reached just using the Internet test scores."*

Test Users

[i] The proposed South African guidelines omit reference to test-takers making informed decisions about the appropriateness of the test, referring instead to obtaining informed consent. It must be kept in mind that in selection situations consent is implied by the test-taker's participation in the process; it is recommended that the ITC wording for this point be retained:

"Assess the appropriateness of the content and technical adequacy of CBT/Internet testing relative to alternative testing methods for each client.

Inform test-takers of the purpose of the testing so they are able to make an informed decision on the appropriateness of the test for their situation.”

[iii] An addition to the ITC point is “Ensure that these policies are not in conflict with legislation and professional regulations”. While ATP SA does not contest this point, determining whether a policy is in conflict with legislation does depend on the *interpretation* of the law (Refer to the legal opinion that will be forwarded to the Board as soon as it is available).

[iv] Reference to verifying test-takers know how to interact with an Internet testing system should still include ‘use of access passwords’ as an illustrative example:

*“[iv] Verify that test-takers know how to interact with an Internet testing system (e.g., basic browser operation, **use of access passwords**).”*

The proposed South African guidelines remove reference to test users informing test-takers of the limitations of the Internet test, as well as providing contact points for the test-takers. In view of the general viewpoints of ATP SA, it is recommended that these two points should be included as Point vii and viii under the Test User section:

“[vii] Inform test-takers of the limitations of the Internet test in terms of the professional relationship expected from this medium.

[viii] For Internet testing, provide a contact point (e.g., email or phone) for those who do not understand the purpose of the test.”

4.2 Consider the psychometric qualities of the CBT/Internet test

Test Publishers

[iii] An addition to the ITC point is “, and that have been evaluated and classified by the Professional Board for Psychology.” This point is dependent on the outcome of the proposed reclassification of psychological tests and ATP SA’s previous submission in this regard. The suggestion made by ATP SA – implementing the EFPA test review model – would

facilitate the sharing of information about tests and allow users to easily verify the scientific base and acceptability of the psychometric property of tests.

Test Users

- [iii] "Use only tests with documented properties, and..." read originally in the ITC guidelines, "Those with documented evidence..."
- [iv] An addition to the ITC point is "that have been evaluated and classified by the Professional Board for Psychology." This point is dependent on the outcome of the proposed reclassification of psychological tests and ATP SA's previous submission in this regard. Also refer to the comments made regarding Section 2.1.

4.3 Where the CBT/Internet test has been developed from a paper and pencil version, ensure that there is evidence of equivalence

Test Developers

- [iii] The proposed South African guidelines omit the example of conditions the intended target population will experience provided by the ITC, which referred to "unproctored or unstandardised testing conditions".

Test Publishers

- [iii] As with the previous *Test Developers* point, "unproctored or unstandardised testing conditions" should be included as an illustration:

"If the developer does not provide evidence relating to the use of the test under conditions that represent those that the intended target population will experience (e.g., unproctored, unstandardised testing conditions), additional studies of test equivalence and norming should be conducted."

Test Users

The final point made in this section of the proposed South African guidelines is "Only use the test in those modes of administration for which it has been designed". This is a valid point – critical to best practice – but the proposed document does not provide an example, which would add illustrative information. It is suggested that the original Point iv of the ITC guidelines be retained in its entirety:

"[iv] Only use the test in those modes of administration for which it has been designed (e.g., do not use a test in an unproctored mode when it is specified for use only in proctored modes."

4.4 Score and analyse CBT/Internet testing results accurately

Test Publishers

[iv] Please note that this point contains a typo, and should read "**Stress** to test users..." instead of "ss".

4.5 Interpret results appropriately and provide appropriate feedback

Test Publishers

[i] In this point concerning informing test users of the potential limitations of interpreting results using CBTI, two valuable points regarding the context of completion laid out in the ITC's guidelines have been omitted from the proposed South African guidelines. ATP SA does not agree that the Supervised and Managed modes should be the only modes of administration allowed in South Africa. As such, it is strongly recommended that they be included as Sub-points c and d here:

"c. for open or controlled modes of Internet testing, test-takers may have been tested in non-standardised, unproctored, or variable conditions, whereas score interpretations are based on administration in proctored, standardised conditions;

- d. *some tests are completed in an administration mode that makes it impossible to guarantee the true identity of the test-taker."*

Test Users

- [i] This point begins by introducing potential limitations the test user should be aware of when interpreting CBTI results, but the sub-points have been omitted in the proposed South African guidelines. ATP SA recommends that this point should read:

"[i] When interpreting the CBTI results, be aware of potential limitations, general and specific, to the reports being used. For example:

- o *Score interpretations are based on administration in proctored, standardised conditions and the test has been administered under open or controlled modes and there is no evidence provided to support the validity of the report under such conditions.*
- o *Tests are completed in an administration mode that makes it impossible to guarantee the true identity of the test-taker.*
- o *Tests alone, however administered, may not provide a complete assessment of an individual, as other confirmatory or ancillary information is not considered."*

- [ii] This point should fit under the potential limitations listed in Point i (refer to the third bullet point above).

- [vi] There is a typo omitting some letters: the beginning of the point should read "**Where** possible, edit..."

- [x] This point has been altered to address providing feedback using computer-based test interpretation, where the original point in the ITC guidelines refers to providing feedback using the Internet:

"Section 30(9) Take account of ethical issues surrounding the provision of feedback using the Internet (e.g. the difficulty of knowing the effect of providing negative feedback to a test-taker, the lack of knowledge of the emotional state of the test-taker,

or the difficulty of providing immediate support to a test-taker when feedback has a negative impact). Where appropriate, feedback should include directions on how to access support and other information.”

4.6 Consider equality of access for all groups

Test Developers

[i] It is suggested that the guidelines should distinguish between bias and fairness – the latter being a more subjective socio-political and contextual evaluation that does not necessarily include a scientific ‘psychometric’ evaluation. It is interesting to note that a South African study found applicants’ perceptions of the use of unproctored, unsupervised online assessments in the selection process to be fair, and generally chose online as their preferred medium of application. This is in line with international findings (Refer to the study presented in Appendix E, which covers applicant perceptions of fairness).

[iv] Using the ITC guidelines as reference, Points v and vi in the proposed South African guidelines should be presented as sub-points under Point iv, concerning guidelines when using tests internationally:

“[iv] For tests that are to be used internationally:

- o avoid the use of language, drawings, content, graphics (etc.,) that are country or culture specific.*
- o where culture specific tests may be more suitable than culturally-neutral ones, ensure that there is construct equivalence across the different forms.”*

[vii] An addition to the ITC point is “and/or other nationally accepted guidelines for test adaptation and validation”. Reference can be made to the Society for Industrial & Organisational Psychology of South Africa’s (SIOPSA) 2005 *Guidelines for the validation and use of assessment procedures for the workplace.*

5 Provide Appropriate Levels of Control over CBT and Internet Testing

5.1 Detail the level of control over the test conditions

Test Users

- [iii] The proposed guidelines state "When testing via the Internet, provide instructions to test-takers that specify the best methods of taking the test." ATP SA fully agrees with this as it places the responsibility on the Psychologist to ensure acceptable test administration practice. This could include unsupervised internet based administration. This should however always happen under the control of the Psychologist.

5.2 Detail the appropriate control over the supervision of the testing

Test Developers

- [ii] This point – defining the managed mode – would be more appropriate placed as Point d of the first point in Section 5.2, where the different levels of supervision required are defined, i.e.:

- "[i] *Document the level of supervision required for the CBT/Internet test.*
- a. Open mode – No direct human supervision required.*
 - b. Controlled mode – Although direct human supervision is not required, the test is made available only to known test-takers.*
 - c. Supervised mode – Test users are required to log on a candidate and confirm that the testing was administered and completed correctly.*
 - d. Managed mode – A high level of human supervision and control over test-taking conditions is required (as in a dedicated test centre)."*

The proposed guidelines state "NOTE that in South Africa only Supervised and Managed modes are allowed, and testing sessions need to be supervised by persons registered with the HPCSA to administer the tests being used."

ATP SA is in disagreement with this statement: it does not align South Africa to international best practices. This effectively restricts Internet testing to a test centre approach where the candidate is sat down in front of a computer and a supervisor logs him/her in, keeping an eye on the whole procedure.

The Board is referred to ATP SA's previous submission concerning the evaluation and categorisation of tests. Not all tests require the same supervision.

Still concerning the level of supervision required, Section 37(2) of the ITC guidelines provides an important point that has been omitted from the proposed South African guidelines, namely:

"[ii] Provide documentation for the testing scenarios for which the CBT/Internet test has been designed."

Test Publishers

[i] An addition to the ITC point is "Note that in South Africa only Supervised and Managed modes are allowed, and testing sessions need to be supervised by persons registered with the HPCSA to administer the tests being used." Again, this point goes against current practice in Internet-based testing in South Africa and around the world and ATP SA stands in opposition to it.

[ii] The final point made in this section of the proposed South African guidelines is "Specify and restrict that use of specific CBT/Internet tests for particular testing scenarios". The proposed document does not provide an example, which would add illustrative information. It is suggested that the original point of the ITC guidelines be retained in its entirety:

*"[ii] Specify and restrict that use of specific CBT/Internet tests for particular testing scenarios. **For example, psychometric tests for use in post-sift selection testing and/or post-hire assessment normally would not be available in open mode.**"*

Test Users

- [i] An addition to the ITC point is "Note that in South Africa only Supervised and Managed modes are allowed, and testing sessions need to be supervised by persons registered with the HPCSA to administer the tests being used." The previous comments made above refer.
- [ii] An addition to the ITC point is ", and in compliance with South African legislation and professional regulations". As ATP SA has already pointed out earlier in this document, determining compliance with legislation is subject to correct interpretation of the said legislation (Refer to the legal opinion to be forwarded to the Board).

5.3 Give due consideration to controlling prior practice and item exposure

Test Developers

The wording of a number of Points i to vi in the proposed South African guidelines has changed somewhat from the original version in the ITC guidelines, with the result that the meaning changes or pertinent information is omitted. It is suggested that the points be presented as in the ITC guidelines, namely:

- "[i] For high-stakes Internet-based tests, use software that tries to equate item exposure rates for items drawn from item banks.*
- "[ii] Limit pilot testing of items on live tests, to minimise unnecessary exposure.*
- "[iii] Make sure item banks are sufficiently large to permit making multiple parallel forms secure and to manage item exposure rates in adaptive testing.*
- "[iv] When parallel forms of a test are created, undertake appropriate psychometric analysis to document their equivalence.*
- "[v] Contemplate delivery strategies that deter memorisation of test content (e.g. by generation of unique tests for each candidate from item banks; or by use of computer adaptive testing).*

"[vi] Control exposure of fixed forms in geographies where cheating is more prevalent by restricted administration to supervised or managed modes."

Test Publishers

- [i] Delete the reference to 'high-stakes' in this point, as 'high-stakes' generally refers to a situation where tests are being used for important decisions such as is the case in selection. Maximum performance tests – typically ability tests where there are correct and incorrect answers – are not solely used online for selection purposes.

Test Users

- [i] Edit this point to include reference to multiple forms of a test, i.e.:

*"Document for test-takers the equivalence of parallel **or multiple** forms of a test."*

- [iii] "Where appropriate, provide test-takers with practice without compromising the security of the **actual test items themselves.**"

5.4 Give due consideration to control over test-takers authenticity and cheating

Test Publishers

- [ii] This point is confusing in the light of the restrictions the Board proposes on Internet testing. Taking into account ATP SA's overall recommendations for CBT/Internet testing in South Africa, it is suggested that this point is presented as it is in the ITC guidelines:

*"[ii] For moderate or high stakes assessment **involving multiple stages**, provide information on how test users can reduce the risk of test-taker cheating (e.g., having another person to take the test as a proxy). **Where an assessment is carried out in open or***

controlled mode, checks against cheating can be carried out by requiring the test-taker to undertake a subsequent validation assessment in proctored conditions (i.e. supervised or managed conditions) and a comparison of scores made.”

Test Users

- [i] This point refers to ensuring test-takers provide the appropriate level of authentication before testing begins. Presenting government approved picture identification is only appropriate in the supervised and managed testing conditions, so the second reference of this point to ‘controlled testing conditions’ needs to be adjusted:

*“[i] Ensure test-takers provide the appropriate level of authentication before testing begins. Remind test-takers (in the Controlled mode) of the need to obtain a password and username to access the test. In **supervised and managed** testing conditions, test-takers should be required to provide authentic, government approved picture identification.”*

Points iv to vi relate to ways in which scores obtained in through a controlled mode should be confirmed. As such, a more coherent presentation of the information would be to include these three points as sub-points under Point iii:

- [iii] For moderate and high stakes assessment (e.g., job recruitment and selection), where individuals are permitted to take a test in controlled mode (i.e. at their convenience in non-secure locations), those obtaining qualifying scores should be required to take a supervised test to confirm their scores.*
- Procedures should be used to check whether the test-taker’s original responses are consistent with the responses from the confirmation test.*
 - Test-takers should be informed in advance of these procedures and asked to confirm that they will complete the tests according to instructions given (e.g. not seek assistance, not collude with others etc).*
 - This agreement may be represented in the form of an explicit honesty policy which the test-taker is required to accept.”*

6 Make Appropriate Provision for Security and Safeguarding Privacy in CBT and Internet Testing

6.1 Take account of the security of test materials

Test Publishers

- [i] The proposed guidelines state "Protect sensitive features of the test from illegitimate disclosure. For Internet testing, all important intellectual property (e.g., scoring rules, norms, interpretation algorithms) associated with a test should remain on the host server. Only test items and the outputs from report generators usually should appear on the test user's or test-taker's screens." Please refer to our comments on Section 2.1; this is an incredibly important point for international and national test publishers.

Appendix A: ATP SA Response

Re-classification of Psychological tests on the current list of classified tests
(submitted to the HPCSA on 30 June 2006)



"ATP SA
Reclassification Proce



"Appendix A Global
Trends.pdf"



"Appendix B Grading
of Psychometric Test:



"Appendix C EFPA
Test Review Form.pd



"Appendix D Test
Administrator Compe



"Appendix E BPS
Test Review Process.

Appendix B:

**Mode of Administration and the Stability of the OPQ32n:
Comparing Internet (controlled) and Paper-and-pencil
(supervised) Administration**

Based on a Masters Dissertation by Gustav Holtzhausen at the University of Pretoria (2004)

Mode of Administration and the Stability of the OPQ32n: Comparing Internet (controlled) and Paper-and-pencil (supervised) Administration

*Based on a Masters Dissertation by Gustav Holtzhausen
at the University of Pretoria (2004)*

Due to the benefits offered by Internet-based testing, more and more test publishers are now using the Internet to deliver questionnaires to clients. Test users and those who are considering or using the Internet as a delivery platform are now facing new ethical, fairness and best practice issues. It is therefore becoming increasingly important to study the effects of the Internet on the fairness and other properties of occupational assessments in South Africa. This study examined the psychometric properties of the OPQ32n when comparing web-based controlled administration with paper-and-pencil supervised data in this regard, to establish whether the controlled web-based administration of the OPQ32n impacts on the stability of the questionnaire.

Two sample groups of managers completed the OPQ32n. The first sample consisted of managers from different industry sectors and with various job functions, which completed the paper-and-pencil version of the OPQ32n in standard supervised test conditions. The second sample consisted of managers in the mining industry across various job functions, completing the OPQ32n on the Internet. To ensure that biographical information does not act as a moderator variable, the samples were compared for significant differences in gender, age and ethnic origin. No significant differences were reported, thus the sample groups were considered to be homogeneous.

The analysis of stability involved comparing the supervised paper-and-pencil sample with the controlled web-based administration sample in three ways. Firstly, the internal consistency as measured by an alpha coefficient was examined and it was found that controlled web-based administration does not compromise scale reliabilities. Secondly, the effect sizes (d -statistic) of the two administrations were compared and only small differences were found in the scores of the two samples. The highest d -statistic is still a small d of 0.35 indicating that there is an overlap of between 78.7% and almost 100% of the constructs between the supervised paper-and-pencil and the controlled web-based groups. Finally, a comparison of covariance structures was carried out using structural equation modelling with EQS. The analysis suggests that the pattern of relationships between scales can be considered the same for both controlled web-based and supervised paper-and-pencil groups as the model fits equally well for both samples, across and within different modes of administration.

The results indicate that the controlled web-based sample and the supervised paper-and-pencil sample have comparable psychometric properties in terms of reliability and covariance structures. This implies that there is no distortion to the instrument itself.

Appendix C:

**Mode of Administration and the Stability of the OPQ32i: Comparing
Internet (controlled) and Paper-and-pencil (supervised)
Administration**

Appeared in SHL Newslines December 2005

Mode of Administration and the Stability of the OPQ32i: Comparing Internet (controlled) and Paper-and-pencil (supervised) Administration

Appeared in SHL Newslines December 2005

The development of computers in the twentieth century has moved the focus from the traditional paper-and-pencil tests to computer-based testing. Computerisation has strongly influenced psychological research and practice. The Internet has now become part of our everyday lives and opened up new opportunities for advancing the science and technology of testing. Advantages of Internet testing include greater efficiency and lower costs (i.e. no printing costs), a reduction in the possibility of missing data (i.e. applicants who do not complete the full questionnaire), the capacity for applicants to be tested from diverse locations and at different times, and the potential for immediate scoring.

In recent years many test publishers have developed personality questionnaires, surveys, inventories and cognitive ability tests to be administered on the Internet. Unfortunately, this has also included tests that are not scientifically validated. Test users and those who are considering or using the Internet as a delivery platform are now facing new ethical, fairness and best practice issues. It is therefore becoming increasingly important to study the effects of the Internet on the fairness and other properties of occupational assessments.

The International Test Commission (ITC) has responded to the rapid growth in computer- and Internet-based testing by developing the International Guidelines on Computer-based and Internet Delivered Testing. The ITC Guidelines addresses issues of good practice in computer-based testing and in particular testing over the Internet. It is applicable to a number of modes of supervision and to a number of testing scenarios. The ITC identified four modes of test administration:

Open mode: Anyone can access and complete the test without supervision; no user identification	Controlled mode: No direct supervision, but access to the tests is via logon name and password only
Supervised mode: Certain degree of supervision; supervisor handles the login process and can verify the test taker's identification	Managed mode: High level of supervision and the test-taking environment is controlled

Initial studies indicate that measurement equivalence between web-based and paper-and-pencil tests is generally established. It is important to investigate the construct equivalence of Internet-based and paper-and-pencil non-cognitive measures in real applicant contexts. The current study focuses on how people behave under real rather than laboratory conditions, but also investigates the effect of supervised versus unsupervised settings.

Two different South African high stakes samples (samples of individuals tested for selection) were included in this study. The first sample (proctored paper-and-pencil) was drawn from an existing database containing applicants tested for various positions in different industry sectors. The second sample (unproctored controlled Internet-based) involved students in their final year of study or who had already completed a degree that applied for various positions at a financial institution. To ensure that moderator variables do not affect the results, an approach was adopted that identified matched samples. No statistically significant differences between the biographical variables of the two groups were found and the sample groups were considered to be homogeneous.

The main objective of this study is to determine equivalence of the OPQ32i administered unproctored on the Internet and proctored with paper-and-pencil in a high-stakes setting. This was achieved by comparing the samples on the mean scores, reliability and analysis of covariance structures. The first analysis examines the mean differences between the two samples by looking at effect size. Cohen (1988) has developed a convention for interpreting effect size and suggests that an effect size or d statistic of 0.20 can be considered small, 0.50 considered medium and 0.80 considered large. The effect sizes of the differences in means between the online unproctored and the paper-and-pencil supervised administration for the current sample ranged from very small to medium.

The second analysis explores the internal consistency of the instruments in supervised and unsupervised conditions. Generally, alpha coefficients between 0.60 and 0.80 are considered reasonable for personality instruments. Alpha coefficients for the supervised paper-and-pencil sample ranged from 0.58 to 0.85 with an average reliability of 0.73 (SEM=2.17). For the Internet-based sample ranged from 0.57 to 0.91 with an average reliability of 0.74 (SEM=2.04). The average reliability for the two samples is very similar indicating that administration mode does not compromise scale reliability.

The final analysis examined the effect of the unproctored Internet-based administration on the pattern of scale intercorrelations using EQS. The scale covariance structures of the two samples were directly compared with prior removal of one scale from the correlation matrix, so that the degrees of freedom become equal to the number of scales. The model tested was that all the correlation matrices were samples from the same population. This analysis suggests that the relationships between scales can be considered the same for both traditional supervised paper-and-pencil and controlled Internet-based testing.

The results indicate that the paper-and-pencil and controlled Internet-based samples have comparable psychometric properties in terms of reliability and covariance structures indicating no distortion to the instrument itself. This South African study supports the available accumulated global evidence that personality

questionnaires can be used for high-stakes assessments via the Internet in a controlled administration mode.

Appendix D:

Testing on the Internet: A validity study of ability tests using Item Response Theory (IRT)

Based on a Masters Dissertation by Esther Venter at the University of Pretoria (2005)

Testing on the Internet: A validity study of ability tests using Item Response Theory (IRT)

*Based on a Masters Dissertation by Esther Venter
at the University of Pretoria (2005)*

A worldwide concern erupted amongst psychologists and related professionals regarding the psychometric properties of tests being published on the Internet (Buchanan, 2003). The purpose of this study, then, was to evaluate the validity of Item Response Theory (IRT) based ability tests when administered on the Internet without supervision. In particular, the study determined the relationship between verbal and numerical ability tests from the Ability Screening Online (ASO) and the Managerial and Graduate Item Bank (MGIB) test batteries. The relationship between these tests would contribute to our understanding of the ASO tests' construct validity (convergent evidence) and would provide useful insight into what the tests are measuring (or not measuring) (SIOPSA, 2005; Saville & Holdsworth (Pty) Ltd., 2002).

The study was conducted using a sample of 124 applicants for a Graduate Development Programme (GDP) that completed both online and paper-based ability tests. The online tests were completed without any supervision and the results were compared with those of the paper-based test results completed under supervised assessment conditions.

Statistical analysis conducted on the data indicated that the online verbal (OVT) and numerical (ONT) tests could be used as predictors of performance on the paper-based verbal (NMG3) and numerical (VMG3) tests. The difference between the average scores for the white group compared with the black group on both the verbal and numerical tests were significant with large effect sizes. This is in line with the differences found on the Management and Graduate Item Bank (MGIB) verbal and numerical tests, which typically range between 0.5 and 0.75 of a SD difference, with white groups consistently outperforming ethnic minority groups (Saville & Holdsworth (Pty) Ltd., 1991; Saville & Holdsworth (Pty) Ltd., 2003a; Saville & Holdsworth (Pty) Ltd., 2003b).

The nature of the relationship between the online and paper-based test scores as well as the results of the leverage point analysis suggest that the influence of cheating on the online scores is likely to be low. This could be due to the "honesty contract" accepted by all applicants or the notification to applicants of a subsequent assessment.

Appendix E:

Applicant Perceptions of the Fairness of Online Assessment in the Selection Process

Based on a Masters Dissertation by Kim Dowdeswell at the University of Pretoria (2006)

Applicant Perceptions of the Fairness of Online Assessment in the Selection Process

*Based on a Masters Dissertation by Kim Dowdeswell
at the University of Pretoria (2006)*

With the growing use of the Internet in the business world, an increasing number of companies are improving their recruitment process by going online. The inclusion of online assessments is a natural outgrowth of this trend made possible by the ongoing advancements in technology and testing theory. However, despite this and the ease with which they may be administered, concerns over online assessments still exist. International best practice requires that, regardless of the mode of administration, the fundamental principles of testing still remain.

In South Africa, in addition to best practice, the law requires psychometric testing to be valid, reliable and fair. Previous South African studies conducted have provided support for online assessments in terms of the instrument's validity (Holtzhausen, 2004; Kriek & Joubert, in progress) and reliability (SHL, 2005). The aim of this study is to determine whether applicants perceive the use of online assessments in a selection process to be fair, thereby adding to the body of evidence supporting online assessment in South Africa. To achieve the goals of this study an online survey questionnaire was administered to an applicant pool following the applicants' completion of online assessments as part of a selection process. The survey data was analysed to determine how the applicants perceived the online assessments in general, whether they considered them to be fair, and the influence of procedural issues as well as the assessments' confidentiality and perceived job-relatedness.

The results found are positive: the majority of respondents are in favour of the online assessments. Additionally, the results suggest that no difference exists in fairness perceptions between Blacks and Whites. Equality of access is also a non-issue, with 82 percent of the sample hailing from previously disadvantaged groups. The few concerns raised are covered by the International Test Commission's (2005) *International guidelines on computer-based and Internet delivered testing*, suggesting that the situation in South Africa is consistent with the international arena. Overall, online assessments appear to be perceived as fair by South African graduate applicants. The applicants' generally positive view of the online assessments may be indicative that in future, in line with international trends, South African applicants will be increasingly accepting of completing assessments online.