Psychological Assessment in South Africa

Published by Wits University Press

Psychological Assessment in South Africa: Research and applications.
Project MUSE. muse.jhu.edu/book/60601.

For additional information about this book
https://muse.jhu.edu/book/60601

For content related to this chapter
https://muse.jhu.edu/related_content?type=book&id=2172654
N. Tredoux

Test users need to be aware of the complexities involved in the use of computerised tests in South Africa. Firstly, it is important to understand that not all computerised tests are the same, and that they differ greatly in their sophistication. This has implications for the professional decisions we make regarding the use of particular tests. This chapter begins with a discussion of the different ways in which computerised testing can be understood. We shall also review the historical background of computerised testing in this country with regard to both the technical aspects and the regulatory framework of psychological tests, in order to help users make informed decisions regarding the use of computerised tests.

What does ‘computerised testing’ mean?

Not all ‘computerised’ tests rely on computer technology to the same degree, or were designed and programmed for implementation on computers with the same level of skill and sophistication. Several aspects of testing can be computerised, such as the administration of the test, scoring and norming, and the generation of narrative reports or data summaries. Sometimes, for a particular test, not all these aspects are implemented to the same degree of sophistication, or at all. Test administration systems also vary in how they employ computer technology. If a test is delivered ‘online’, for instance, it means that the program that controls the test administration is located on a server somewhere on the internet, and not on the computer used for administration. Online, or internet-based, administration creates the possibility of administering a test in an unsupervised way, at the respondent’s convenience. However, it is quite possible to administer an online test with some degree of control, provided that the necessary technology is in place and is properly utilised.

Scoring and norming

The simplest form of computerisation in psychometrics is the use of computer software to score multiple-choice tests that have been administered using pencil and
paper. Data input can be manual, using a spreadsheet or similar software, or even software especially written for the test in question. It can also be mechanised, or purely electronic. Early forms of mechanised data acquisition involved special cards, with the respondent punching a hole in the card to indicate his or her answer. Optical mark readers have been the predominant technology for mechanical input of test responses for the past few decades. These machines require the test to be answered on specially printed answer sheets. These sheets can sometimes be difficult to read and complete, which can interfere with the reliability of the measures. They also sometimes require the use of a special pencil, and errors in answering are difficult to correct. The scoring key is also fed into the computer, either as a separate data file or coded into the scoring software. The same applies to norm data. The software then scores the responses, relates the raw scores to the norms and can print out raw and normed scores. Often this is done in the form of a profile.

The advantage of computerised or computer-assisted scoring lies in the reduction of errors, and in the considerable time-saving that can be achieved. It is especially valuable for questionnaires that measure multiple dimensions. The following types of scoring and norming errors that occur with pencil-and-paper scoring are eliminated:

- using the wrong scoring mask;
- skewed orientation and/or misalignment of the scoring mask;
- miscounting of item endorsements;
- errors in transcribing scores;
- errors in looking up standard scores in norm tables; and
- incorrect positioning of scores on profile sheets.

Even when computers are used to score tests, errors can still occur, and users should be aware of the following sources of error that can occur with the use of computer technology:

- When doing manual response capturing, the capturer can get out of step in typing in the responses, and end up typing the wrong response against a particular question.
- Whether using optical mark readers or manual capturing, the wrong test for a particular answer key can be chosen in the software.
- Errors can also occur when specifying the norm groups in the software.
- Errors in capturing biographical details or creating new database records for respondents can result in inaccurate reporting and corrupted data.
- Answer sheets that are designed for optical mark readers are typically more difficult to complete, and respondents can make errors that can cause their protocols to be unscorable or inaccurate. Test users need to be vigilant for this during test administration.

Computer-generated reports

Computers can make a more sophisticated and potentially valuable contribution to the assessment process by generating a narrative report from normed scores.
The quality and style of such computer-generated narrative reports vary widely depending on the system used and the particular reporting program. The best of them are difficult to distinguish from reports written by expert psychologists, and are even able to integrate results from different tests. The most elementary computer-generated reports do little more than give a description of the constructs that have been measured, with the level of score the respondent attained. The use and interpretation of computer-generated reports is discussed in more detail later in this chapter.

Test administration

Computers can be programmed to present test instructions, examples and test items. In most cases this is done by means of text displayed on the computer monitor. Sometimes voice recording or synthesised speech can be used. Highly sophisticated test administration systems can handle a multitude of response formats, and some can adapt the items to the respondent’s level of ability while the test is being administered. The more interactive and sophisticated the computerised test administration system, the more expensive it is to develop. It must be recognised that computerised tests force respondents into a more structured manner of answering compared to pencil-and-paper tests where, for instance, it is much easier to refer back to earlier answers, or use certain answering and checking strategies (Bugbee & Bernt, 1990). Not all candidates react equally positively to computerised test administration (Moe & Johnson, 1988), and there may be systematic differences between population groups in terms of how they experience being tested on a computer (Legg & Buhr, 1992).

The speed of technological progress

Computers decrease in size and price and improve in speed, storage capacity and visual display capability at an astonishing rate. The resolution and flicker of computer screens used to be a concern from an ergonomic point of view, but with the quality of computer displays available at the time of writing, those concerns appear to have become irrelevant. Rapid advances continue in terms of technologies such as voice synthesis and recognition, text and handwriting recognition, as well as the processing and analysis of text. Future testing systems will not need to be limited to the scoring of multiple-choice responses, or even need to receive responses through mouse-clicks, touch screens or keyboards. Open-ended questions, written responses, gestures and spoken responses can all be processed by computers already, even if these capabilities have not yet been incorporated into psychological testing systems. These advances could mean that with sufficient ingenuity and investment, the computer literacy of respondents as a limiting factor for psychological testing could be almost eliminated. Rapid technological advances also mean that test developers and even test users need to be proactive in staying technologically
Using computerised and internet-based testing in South Africa

aver, in order to evaluate the significance and implications for psychometrics of new developments.

The pervasiveness of computer technology

Computer technology used to be limited to organisations such as universities, government departments, research institutes and large corporations. When computers were rare and expensive, they were also relatively primitive. Now many individuals have computers of their own. People who do not have access to their own computer can use public computers that can be found in schools, libraries, community centres and internet cafés. Domestic appliances, motor cars and cellphones contain processors that would, in the past, have been considered relatively powerful in terms of computing capability. Cell phones, electronic game consoles, music players and other electronic devices can have internet access. These devices can be highly portable and compact. This means that one cannot assume that when a test is delivered over the internet, it will be completed in an office or quiet home environment, on a computer with a screen of a certain size. This has implications for the standardisation of test conditions when internet-based test administration is used.

Unequal access to computer technology

Even though computer technology is becoming very widespread and inexpensive, it is still mostly available only to fairly affluent people, in relatively urbanised environments (Technology Access Foundation, 2010). The very poor, and people in rural areas, have probably not been exposed to computer technology to the same degree as affluent urban populations, if indeed they have had any exposure to computers at all (Sutton, 1991). Some people have an aversion to computers that may prevent them from becoming comfortable with them, even if they have the opportunity to accustom themselves (Brosnan, 1998). Even with efforts to familiarise respondents with the devices to be used for testing, the fact that some may have had access to similar devices before and others not creates an inequality which raises concern in terms of bias and fairness. It is important that assessment practitioners be observant when testing respondents on computers. They should be alert for respondents who show obvious signs of discomfort with the medium of administration, and watch out for people who struggle with the mouse, keyboard or whatever means is used to enter the test responses. It is helpful to have a nonthreatening computer-based exercise to introduce respondents to the computer. Good rapport and handling of questions during the instructions and early phase of testing are very important. If necessary, an alternative means of assessment should be used if respondents remain very uncomfortable with the computer. Of course, it is very difficult to deal with these issues if the test is administered without supervision.
Computerised testing in South Africa

To understand the current controversy surrounding computerised testing in South Africa, it is important to be aware of the historical background: South Africa’s innovative leadership in computerised testing experienced a setback when concerns were raised about the fairness of psychometric testing. The regulation of psychometric tests, whether pencil-and-paper-based or computerised, is still in dispute.

Early computer-based developments

During the late 1970s and early 1980s, researchers at the National Institute for Personnel Research (NIPR) did pioneering work in computerised testing. A computer-based system to administer tests of intellectual ability was developed on a Varian minicomputer. Subsequently, an extremely ambitious system was written to administer a wide range of psychological tests over a wide area network using touch screens. This testing system adapted the Control Data ‘Plato’ system (Plato Learning, 2010), which was designed for education and training, for psychological testing. It had many groundbreaking features, such as the ability to recover from system interruptions or power failures, resuming where the respondent had left off with all test response data intact.

The NIPR Plato-based testing system allowed the psychologist who controlled it to set up batteries of tests for administration in other cities. Respondents would not necessarily complete the tests under professional supervision, because at the time psychologists were allowed to delegate any psychological action to an unregistered person, provided these actions were supervised (Department of Health, 1977). The Plato-based testing system included its own training module for familiarising respondents with the computer, and all respondents completed this. Respondents who could not answer the example items correctly after a number of attempts were not allowed to complete the rest of the tests. The test responses were routed across a wide area network and processed centrally on a mainframe computer. The system included programs for calculating norms and reliabilities. The Plato-based testing system was implemented at some universities, some state-owned enterprises and the NIPR’s own assessment services. Most of the tests programmed for this system were conversions of existing pencil-and-paper tests. There were, however, some instruments developed specifically to take advantage of the computer system’s interactive capabilities. Among these were a simulation exercise called ‘Maze’ designed to assess managerial decision-making (Tredoux, 1985), and a comprehension test that involved the reordering of sentence segments.

These early developments took place during the apartheid era. At this time, different race groups did not compete for the same jobs. The profession of psychology was also relatively young in South Africa, and there was no official, enforceable code of ethical conduct for psychology practitioners. Most of the respondents who were tested on the system were white; therefore cross-cultural studies were not undertaken on data collected on this system. It was not expected that any of the respondents would be computer-literate, and therefore group differences in computer literacy were not considered a factor that might potentially make the tests unfair.
Sanctions and the move to microcomputers
As a result of political and economic sanctions, Control Data, the company that supplied the Plato system and the hardware on which it ran, withdrew from South Africa. Eventually it was no longer feasible to maintain the Plato-based testing system and it fell into disuse. Soon afterwards, microcomputers became available, despite continuing sanctions. In the mid-1980s the NIPR, by then part of the Human Sciences Research Council (HSRC), developed a comprehensive microcomputer-based testing system called PsiTest. Another system, called Siegmund, was also developed at the HSRC. Both these systems were used fairly widely. The PsiTest system had optional enhancements for managing testing sessions, utilising the capabilities of microcomputer networks. A simulator-type system to test vehicle drivers was also developed at the NIPR, using an Apple computer. A clerical office work simulation for assessment was developed on the Microsoft Windows platform. The HSRC also did research on, and acted as a supplier for, the Austrian-developed Vienna Testing System, which specialised largely in computerised psychomotor testing (Schuhfried Gmbh, 2010).

The HSRC’s microcomputer-based testing systems brought computer-based testing within reach of psychologists in private practice, as well as larger organisations. The systems were sold only to psychologists, and since the internet was not yet functional the tests were always administered under supervision, even if the supervision was not done by a registered person. A registered psychologist was always professionally responsible for the testing session.

Backlash against testing
During the 1980s, psychometric testing in South Africa became subject to increasingly strong criticism. The HSRC was the statutory organisation responsible for research and development in psychometric technology and was ill-prepared to answer the criticisms that it had not paid sufficient attention to bias and fairness (Taylor, 1987). Insufficient cross-cultural research had been done on the HSRC tests. This was partly due to the fact that the HSRC had developed separate tests for different race groups, and partly due to the historical situation that black and white applicants had not been competing for the same positions, and hence had not been tested on the same tests, leading to a lack of comparative data. Around this time, people (including psychologists) were turning against psychological testing, questioning its fairness and even its legality (Taylor & Radford, 1986), which inhibited further data collection. The HSRC started paying explicit attention to the bias and fairness of the tests that it was supplying (Owen, 1989a; 1989b). The South African Personality Questionnaire, in particular, was examined for cultural bias and found wanting (Taylor & Boeyens, 1991). Amid political unrest and eventual changes in government, state-funded research and development in psychometrics were curtailed. Early drafts of the Employment Equity Act banned psychometric testing for employment purposes (Employment Equity Bill No. 60, 1998). When the final version of this Act, the Employment Equity Act No. 55 of 1998, was eventually promulgated, allowing psychometric testing subject to the tests meeting technical psychometric and fairness requirements, many experts in psychometrics had already left the country or joined the private sector.
Computerised testing becomes highly commercialised

Even before the HSRC reduced its involvement in psychometrics and computerised testing in the early 1990s, some publishers in the private sector had already entered the field of computerised testing. One of the early South African initiatives was a report-writing system for the Sixteen Personality Factor Questionnaire (16PF), developed by Professor Dan Steyn. The computerised testing systems that had been supported by the HSRC, such as the Vienna System and the HSRC computer-administered testing system (CAT), were taken over by private sector organisations.

Drs Dawie Minnaar and Pieter Erasmus developed the Potential Index Battery (PIB), which later evolved to become ‘Speex’. This set of tests, each measuring a single dimension, gained widespread acceptance in industry. It featured computerised administration and scoring, and the ability to calculate norms from data that had been collected. Also active in the South African market was Thomas International Ltd with the Personal Profile Assessment (PPA), which produced a comprehensive narrative report from a short ipsative questionnaire (Thomas International, 2010).

The Discus system featured a very similar questionnaire to the PPA, also with computer-generated reporting targeting the same dimensions (Axiom Software, no date). Jopie van Rooyen and Partners acquired the South African distribution rights for a large number of tests and questionnaires. Several of these, notably the 16PF and Myers-Briggs Type Indicator (MBTI), soon featured computer-generated reports available through bureau scoring (JvR Group, 2010). M&M Initiatives (2010) developed the Learning Potential Computerised Adaptive Test (LPCAT), based on item response theory. Dr Terry Taylor formed a company called Aprolab and entered the market with two learning potential measures, Transfer, Automatisation and Memory (TRAM) and (Conceptual) Ability, Processing of Information, and Learning (APIL), that featured computer-assisted reporting although they were originally administered using pencil and paper (Aprolab, 2010). Maretha Prinsloo of Cognadev developed the Cognitive Process Profile (CPP), a computer-administered measure of thinking style and potential (Cognadev, 2010).

Saville and Holdsworth Limited (SHL) entered South Africa in the mid-1990s with a suite of computerised products that encompassed job analysis, personality and ability testing. Dr Kobus Neethling developed a range of instruments measuring thinking styles and preferences (Neethling Brain Instruments, 2010). The Neethling brain profiles were, however, never classified as psychological tests. Psytech International products, also including a range of personality and ability tests with narrative reports, were introduced to South Africa in 1994. Since 1998, access to the Psytech tests and testing software has been restricted to registered professionals (Psytech South Africa, 2010).

The PIB, Discus, SHL, Neethling and Thomas International products were sold for use by people who had been trained by the publisher, and who were not necessarily registered psychologists or psychometricians. Some of the publishers – notably, Neethling, SHL and Thomas International – accredited unregistered users to use their instruments independently of the Health Professions Council.
of South Africa (HPCSA) registration system, in line with their organisations’ international practices (Neethling Brain Instruments, 2010; SHL South Africa, no date; Thomas International, 2010). Within approximately a decade, computerised testing in South Africa had changed from a largely state-funded and controlled activity to a highly competitive, highly commercialised industry. There was and still is a lack of consistency between different providers in the restriction of access to computer-delivered psychological tests.

The advent of internet-delivered testing

In the mid-1990s access to the internet started becoming widespread in South Africa. Very soon, numerous local and international tests became available to anyone with internet access. Some of these were from reputable publishers, but several tests that were clearly psychological in nature were made available by people who had no intention of having the tests evaluated or classified, or of limiting access only to people registered with the HPCSA to use tests. Several of these tests were accessible on a ‘pay per use’ basis, whereby any person could pay an amount via credit card and then was given access to the test and the report. The website hosting the test and processing the transaction could be anywhere in the world, thus bypassing regulations in countries that regulate testing strictly. However, even with websites that are clearly South African and that fail to abide by the South African rules regarding psychological tests, the Professional Board for Psychology has not been successful in exercising control. Where the websites are run by unregistered people, the Board is not able to prosecute and needs to refer the matter to the National Prosecuting Authority. This has not yet resulted in any successful prosecutions.

Considering the scale on which unregistered persons use psychological tests via the internet, there have been remarkably few complaints to the Professional Board. The reason could be that respondents are not aware of their rights, or that, as job applicants, they feel too disempowered to take action. Even more remarkable is that psychology professionals have lodged so few professional conduct complaints with the HPCSA about psychologists allowing unregistered persons to access tests via the internet, although the author has been the recipient of numerous informal complaints.

The International Test Commission (ITC) has formulated guidelines for the use of computerised and internet-based testing (ITC, 2005). These guidelines outline the responsibilities of various stakeholders in the testing process and distinguish between different modes of administration, ranging from managed mode (administration in a special facility with a supervisor present) to open mode (unsupervised self-administration open to any person). Between these extremes is controlled mode administration, where the identity of the respondent is verified and other technologies are employed to prevent cheating, but the administration still essentially proceeds unsupervised. The Professional Board for Psychology has not accepted the ITC guidelines for unchanged application in South Africa. A limited version of these guidelines was published, allowing only managed mode, but the HPCSA eventually withdrew it after legal action by test publishers (ATP vs HPCSA, Pretoria High Court, Case No. 4218/07).
Critical evaluation of the advantages of computer-administered testing

Facilitation of research
Well-designed computerised testing systems will store the test responses in a database that is accessible for research, or will have options that enable users to export the data to a format that is compatible with data analysis programs. To protect confidentiality of the respondents, it should be possible to make the data anonymous. Errors of scoring and transcribing data can be eliminated if the test is administered directly on the computer. If a test is in development, items can be trialled on computerised testing systems without the cost of printing pencil-and-paper materials that will need to be discarded afterwards. New, updated versions of tests and norms can be made available quickly and inexpensively over the internet, rather than requiring users to purchase new printed copies. These are enormous advantages for test developers, and also make it more feasible to do the psychometric and cross-cultural research that is required in South Africa for compliance with the Employment Equity Act. Hence, well-designed computerised testing systems can also help to protect the rights of respondents by ensuring that up-to-date norms and item sets are used.

Monitoring of usage
From the point of view of test distributors and developers, the fact that computer systems can count the number of times a particular test is used, and by whom it is used, is an important benefit. Pencil-and-paper tests are very vulnerable to copyright violations which cost test developers a lot of money, thus inhibiting further research and development. Computerised test administration makes it much easier for test developers to ensure that they profit from their efforts.

Standardisation of administration
Insofar as test administration consists of presenting instructions and test items to the respondent, timing the responses and scoring them, computers can do that very well. The computer is dispassionate and objective. It treats everybody in exactly the same way. But is this all that test administration should be in a society where there are numerous obstacles to fair administration in terms of culture, language and computer literacy? Is it professionally justifiable in South Africa to allow respondents to complete tests unsupervised over the internet? An examination of the ethical code and defined competencies regarding test administration, as specified by the Professional Board for Psychology, suggests that this is not the case.

The competencies that a psychology professional should master with regard to test administration include more than the accurate conveying of instructions and test items, timing and scoring. These competencies are spelled out with particular reference to psychometrists, although they apply to other categories of professionals who administer tests as well (Professional Board for Psychology, 2006). Among other things, the psychometrist is supposed to ensure that the environment in which testing takes place is conducive to testing. The psychometrist should
observe the respondents to consider whether they are in a fit state to be tested. There should be a process of building rapport, obtaining informed consent for testing and observation of behaviour during testing. The psychometrist should evaluate whether all the respondents understand the instructions fully, and give special attention when necessary. The psychometrist should also make sure that the respondents do not cheat, obtain help during testing or remove test materials from the test room. The psychometrist should consider whether the use of an interpreter may be required. In a computerised test room, the psychometrist should consider, by observation, whether all the respondents are coping with the computer interface, and whether an alternative mode of testing would not be more appropriate. The psychometrist should also deal with the eventuality of a power failure or equipment malfunction and its effect on the respondents. Much of this professional administration process is aimed at protecting the right of the respondent to be treated fairly. Test administration according to the standards specified by the HPCSA is not just the mechanical process of reading instructions, timing and scoring. Rather it is a professionally compassionate, interactive process of observation and adaptation of the testing process when necessary, to ensure that respondents are tested fairly and not merely uniformly.

The ethical code for psychologists, now incorporated into the Health Professions Act No. 56 of 1974 as a regulation (Department of Health, 2006), warns psychologists to limit their findings appropriately when there has been any deviation from standard testing practice. They are even warned to limit their conclusions when group test administration rather than individual test administration has been done. They are also instructed to limit their conclusions when using any computer-mediated processes. The ethical code further specifically states that psychological assessment must take place in the context of a defined professional relationship. It is difficult to see how this requirement could be met with unsupervised test administration. In many cases there is no personal contact between the person responsible for the assessment project and the respondent completing an unsupervised internet-based test. Respondents usually simply get an email with a link on which they click to bring up the test. The person in charge of the testing process also usually has no control over the time of day when the test is completed, whether testing conditions are adequate and whether some respondents receive help during testing.

It is thus clear that when tests are administered unsupervised over the internet, the rights of the respondents are not as protected as when the tests are administered under the supervision of a psychologist or psychometrist. Some candidates, particularly those with lower levels of literacy and computer experience, may be more disadvantaged than others. Supervised computerised test administration can, however, have considerable advantages – provided it is properly managed. Systems are becoming available for supervising test administration sessions remotely through web cameras and using instant messaging facilities (Psytech International Limited, 2010). Other verification systems, mainly aimed at controlling cheating, involve doing a verification test under supervised conditions after screening applicants based on an unsupervised test (SHL Group, no date).
Technical considerations

In choosing and implementing computerised tests, psychologists and psychometrists should bear a number of factors in mind. These are discussed below.

System stability

If the computer system is unstable due to hardware or software factors, it is not suitable for testing. The computer should meet the minimum system requirements for the testing software. Security systems on the network or computer, such as antivirus software and firewalls, should not obstruct the functioning of the testing software. The system should also be free of viruses and malware that can slow down performance. Any necessary modules that are needed to display item content, such as special software to display video and animation, sound and video drivers, and so forth, should be installed and up to date. If an internet connection is required, it should be up and running, and the connection speed should be adequate. This is particularly important if the nature of the test administration is highly interactive. It may be necessary to take special precautions to ensure that the testing session will not be interrupted by a power failure. It should be part of the preparation for the testing session to ensure that everything works as it should. If anything on the system has changed since it was last used for testing – such as the updating of the operating system or reconfiguration of the firewall – this is particularly important.

Well-designed testing systems should be able to resume with data intact after an unforeseen termination, and to display test items correctly even when the line speed is slow. However, not all testing systems are capable of doing this. The test administrator should be aware of the risks and vulnerabilities of the particular systems that are being used.

Connection speed

Some assessment systems, particularly those that involve video, require a very fast internet connection. Well-designed assessment systems will download the necessary information before commencing the test. However, in areas where the connection is slow, this can take a lot longer than expected and may interfere with the scheduled start of the testing session.

Systems that allow test administration to be observed remotely through web cameras are particularly vulnerable to slow internet connections. Hopefully this situation will improve as South Africa’s connectivity infrastructure is upgraded. In deciding to use systems that rely on fast internet connections, professionals must consider their environment. If testing will take place in rural areas and will connect to the internet through cellular modems, it is necessary to verify whether connectivity will be available in the area and whether it will be possible to connect at full speed. Some internet services are affected by bad weather. If the connection will be slow, extra preparation time may be required to download the items.

Purchasers should bear in mind that systems developed in other countries, where internet speeds are much faster, may never have been tested with
bandwidth as limited as we have in South Africa. Poorly designed internet-based assessment systems will simply fail, slowing down or halting during testing if the connection becomes too slow. They may even lose information. If the program is not written to take account of variations in line speed, it could result in inaccurate timing of tests. This is not conducive to standardised testing procedure, and psychology professionals should verify that this will not happen before committing to using such a system.

Security considerations

Computer-based and internet-delivered tests need to be programmed with a greater awareness of security than normal applications. Aspects that should be considered include access control, security of item content and scoring keys, and security of results. These will now be discussed in more detail.

Access control

As with pencil-and-paper test materials, access to computerised testing systems should be limited to authorised users. Many systems are password-protected. Users should know how to change their passwords, and not use a password that is easily guessed or used by a lot of other people. Some systems permit several levels of access. Some individuals may, for instance, be able to set up test batteries, customise reports and change norms, whereas others would only be allowed to perform limited administrative functions such as data capturing on the system. Professional users should be aware that when they share their password or give another person access to a computerised testing system, they are potentially delegating actions that may be reserved for the profession of psychology. The ethical code specifies that such actions should only be delegated to persons who are competent and appropriately trained to perform them.

Security of item content and scoring keys

It is important to pay attention to the computer file formats in which test items and scoring keys are stored on the computer. Item text and graphics should be encrypted so that they cannot be accessed by unauthorised people using a word processor or graphics program. Internet-delivered tests are particularly vulnerable in this regard. Some internet-delivered tests leave files containing the test items behind in the internet cache, or temporary file folder. These files can then be accessed and saved or printed after the test has been completed. This is a serious breach of security and can compromise the integrity of the test and the assessment process. Test users should be aware of this and make sure that they do not use systems that ‘leak’ confidential information in this way.

There is a possibility that respondents may deliberately try to copy items while doing tests. This risk is much greater when tests are administered unsupervised and respondents may be completing the test in their own homes or offices. Securely designed testing systems do not allow respondents to make screen copies or printouts while the test is in progress. Many internet-based
tests can be copied by screen-capturing or printing the items, and practitioners should avoid tests that have this security vulnerability. Even if the ‘print screen’ key is disabled, some people may try to photograph the items. As a precaution, it is better not to allow respondents to have cell phones accessible while they do computerised tests, since many cell phones incorporate cameras.

**Security of results**
When obtaining informed consent for testing, the limits to confidentiality must be clarified (Department of Health, 2006). In so doing, the respondent agrees to the persons, other than the psychologist or psychometrist, who may have access to his or her test results and the report based on them. The psychologist then has a fiduciary duty to respect these limits and to make sure that nobody else accesses this confidential information. Psychology professionals in South Africa are also obliged to store psychological assessment results securely for five years. Password-protected databases can be a great help in this regard. Databases should, however, be backed up regularly. Backups should be done onto a different physical device than the one where the original records are stored. Removable media are a good solution. Backup copies should be stored securely, again preferably in a different location to the original data.

Computer-generated reports are usually produced as word-processor documents. Security violations can easily occur when these are saved to hard drives or sent through email, or when printouts are distributed within organisations. It is the responsibility of the psychologist or psychometrist to ensure that unauthorised people do not see psychological reports. The identities of the people who are permitted to see the reports should be clarified when obtaining written informed consent from the respondent before testing. Report documents should preferably also be password-protected. Recipients of the reports should be warned not to have printouts lying around in accessible places, such as in in-trays or on desks. All reports should be clearly marked ‘confidential’.

**Human–computer interface considerations**
When a respondent is completing a psychological test using a computer, he or she should not be struggling to deal with the apparatus rather than attending to the test content. The interface elements – keyboard, mouse and screen – should not ‘get in the way’ of the test items. This is more difficult to achieve for people who have had little experience with computers. However, it is also important to make sure that the equipment being used is of a sufficient quality that it does not in itself place the respondent at a disadvantage.

**Display quality**
The image on the computer screen should be crisp, clear, stable and not distorted. If the equipment is relatively new, this should not be a problem. However, older screens can develop problems that may interfere with test administration. If the screen is of different dimensions than the screen for which the item material
was programmed, items can appear distorted if the program was not written to compensate for this. If groups of people are tested at the same time, they should be tested on equipment of comparable quality to avoid creating an unfair situation.

**Keyboard familiarity**

Some respondents have not had the opportunity to learn keyboard skills. This can be the case with people from economically deprived backgrounds, but also with older respondents who may hold senior executive or professional positions, and do not personally use computers at work because they have support staff who do it for them. Well-designed computerised tests should not demand an inappropriate level of computer skill from the respondent, relative to the construct being measured and the purpose of measurement.

**Pointing devices and touch screens**

In most computerised tests, respondents indicate their choice of answer with a pointing device. Pointing devices come in many types, the computer mouse being the most common. However, light-pens, trackballs, joysticks, touch-sensitive screens and touch pads are also used. Usually, the test program does not distinguish between different types of pointing device. It is up to the test administrator to see that the respondents are comfortable using the particular pointing device. In the author's experience, the most acceptable, inexpensive and reliable pointing device is the ordinary computer mouse, preferably the optical type. The touch pads and little mini-joysticks found on some laptop computers are difficult to use for people who are not used to them and they are best avoided for testing purposes.

**Technological literacy**

Psychology professionals who use computerised testing should consider their own technological literacy as well as the technological literacy of the respondents. If the test administrator is not able to cope with the technical demands of the assessment system, the assessment process may be discredited and could place the respondents at a disadvantage. In considering the appropriateness of computerised testing for a given group of respondents, the psychologist or psychometrist should verify that there isn’t a subgroup who is significantly more technologically sophisticated than the rest. It is differences in technological literacy between people who are being assessed for the same purpose that create unfairness, rather than the overall technological sophistication of the group. This is especially true if lack of technological literacy will affect test performance, and if technological literacy is not in itself relevant to the construct being assessed.

**Advancing the state of the art**

Even when discussing the limitations of computerised testing systems, one must be aware of the fact that this is a field of technology that has the potential for very rapid development. With the technology becoming available, it is possible
to create testing systems that are interactive, responsive to individual needs and secure. However, implementing the new advances will require a commitment from developers. Users also have to be discerning, since many of the systems available today are based on outdated technology.

**Item content**
Multimedia item content in computerised testing is not new, but it is becoming easier and less expensive to implement. The potential this offers is not only in making tests more attractive and exciting, but also in accommodating people who are visually impaired. Tests can now be made more interactive, in the sense that the testing system responds to the person who is completing the test and adapts the testing process accordingly. The task set for the respondent can also be made more meaningful than merely choosing an answer from a number of options. The CPP is an example of how the interaction with the computer can be used to externalise mental processes.

**Processing of responses**
To be truly interactive, a testing system needs to process responses while the respondent is completing the items, and not only afterwards. Computerised testing systems can be programmed to take account of response latencies (the time lag between the presentation of the item and the response), the number of errors made and the number of corrections. Computers can monitor and assess learning that takes place during the testing process. With voice recognition, handwriting recognition and sophisticated text-processing capabilities, computers can be programmed to assess the quality of a person’s thinking as well as tally the number of errors made.

**Supporting advances in psychometrics**
With the capabilities mentioned above, computerised testing systems are essential for implementation of tests that go beyond classical psychometrics. With the limitations of classical psychometrics becoming increasingly apparent, it is essential for test developers to embrace new technologies – for instance, tests based on item response theory. Test users must likewise remain up to date with new approaches to psychometrics.

**Adaptive tests**
Using item response theory and related sophisticated algorithms, it is no longer necessary for all respondents doing a particular test to complete the same set of items in the same sequence. Testing systems can adapt the difficulty level of the items to the candidate and thus test them more accurately and economically. This approach also makes it much easier to protect the integrity of the test, because such tests are much more difficult to copy.

**Advanced reporting**
Users of computerised tests have come to expect narrative computer-generated reports, and in many cases are prepared to pay extra for the convenience. High-quality computer-generated narrative reports can appear so credible that the
Using computerised and internet-based testing in South Africa

439

test user could even be tempted to overlook the psychometric limitations of psychological tests and take them at face value.

Computerised narrative reports are not inherently difficult to produce. They can make use of artificial intelligence, but this is not necessary for a good report. Computer-generated reports essentially need to consider all the different possible score combinations and generate text for them. They do require a great deal of work, since text needs to be generated for very large numbers of score permutations. This requires not only programming expertise, but also conscientious application and cooperation from insightful psychologists with expertise in the tests. There are specialised software systems available that assist in the customisation and automation of computer-generated reports. An example is the GeneSys system from Psytech International (Agnew, 2003). It is even possible for a knowledgeable user to program a narrative report using spreadsheet software or word-processing software and a database. This is, however, beyond the competence of most psychology professionals, and they rely on the test vendors to provide them with the reporting technology.

Simple reports that merely report on the scores for a single test are the most common. However, with sufficient effort, expertise and investment, it is possible for computer-generated reports to move beyond reporting into the realm of interpretation. Reports can also be developed that integrate results from a whole battery of tests. Advanced reports can evaluate a person’s test scores within a specified context. For instance, the report can compare different score combinations against the desired profile for a given position, which can be very useful when using tests for selection. The report-writing program can calculate how well a respondent can be expected to perform on specific work-related competencies or in a specific role, and give an explanation of what can be expected from the person, and which of his or her core characteristics give rise to this expectation. Advanced computer-generated reports can guide a manager in overcoming a respondent’s development needs and maximising his or her potential strengths. A very useful type of report is one that generates a follow-up interview schedule that enables the professional to probe or clarify certain measured characteristics. This can act as a means of collecting additional information, and also help to verify or clarify the findings from psychometric tests. Reports such as these enable test users to acquire the ability to use a test in a sophisticated manner very quickly. What used to take years of experience, training and supervision can now be made possible very quickly with the help of a computer.

In some cases, the report-writing program does calculations on scores before interpreting them. These calculations could be estimations of certain dimensions that were not directly measured by the test. They are often called ‘derived scores’. The formulae used to do these calculations are sometimes based on empirical research and sometimes on expert opinion. Both of these sources of information contribute to raising the cost of producing the report program.

It is expected that, due to the cost of refining and customising computer-generated reports, the majority of reports will probably continue to be somewhat generic, rather than adapted to a respondent’s individual circumstances and the
requirements of a particular role. It is usually necessary for a professional to do some editing and contextualising to take account of the respondent’s individual circumstances and the specific context in which the assessment is done. Even sophisticated reports are usually developed for the international market, and it is often necessary for test users to edit the reports to incorporate the South African social, cultural and language considerations afterwards.

Professional control

The issue of professional control over computerised testing in South Africa has been controversial. The Professional Board for Psychology accepted a policy on computerised testing which was largely based on the policy published by the ITC. The South African policy, however, did not allow unsupervised testing and placed an age restriction of 18 on clients who could be serviced through computerised means. Test publishers opposed the policy and forced the Board to withdraw it. The policy has been put before the Board several times and has been accepted unanimously on each occasion. Meanwhile, the Health Professions Act was amended, and regulations allowing unregistered persons to perform psychological acts were repealed (Regulation R993, Health Professions Act, September 14, 2008). Moreover, testing for employment was added to the actions reserved for the profession of psychologist. Thus, in terms of the legal and ethical regulation of testing, the situation is now more strictly controlled than when the South African policy on computerised and internet-delivered testing (HPCSA, 2006) was first accepted. Whether there is a specific policy on computerised testing or not, the legal and ethical restrictions on the use of psychological tests remain.

However, South African professionals should take cognisance of the responsibilities for test users listed in the ITC’s guidelines (ITC, 2005). Using computerised and internet-delivered testing requires a higher level of technological sophistication, a greater awareness of security issues and a responsible concern for the welfare of the respondent. Furthermore, our local regulations require that a psychology professional take personal responsibility for assessment work. This is difficult to do with unsupervised testing, where the test user cannot even be certain that the equipment on which the test will be completed will be adequate for testing purposes. Computers should not be used to mass-produce assessments. We must never lose sight of the fact that we work with individuals who have constitutionally protected rights.

References


