APS Australian Psychological Society

Online psychological testing

APS Tests and Testing Expert Group

August 2018

This resource was developed by Peter Macqueen, Wally Howe and Marian Power as members of the Tests and Testing Expert Group.

Online psychological testing

Table of Contents

1. Background

Prior to the advent of the internet, and online testing, computers were used primarily as "page turners" in order to administer and score paper and pencil tests. Hankes is reported to have developed, in 1946, an analogue computer to score the Strong Vocational Interest Blank (SVIB; Moreland, 1992). Nevertheless, more innovative applications were developed and, for example, early research with work sample assessments, administered via computer, included the use of a simplified landing simulation for use in pilot selection (Bartram, 1987).

As computers and the internet became more widely accepted and used, paradigms emerged to encapsulate modern methods of psychological testing. One of the most supported models is that of Bartram (2001) in which he defines four modes of test administration via the computer or the internet:

a. Open: no conditions; no test taker identification (insecure).

b. Controlled: no supervision, but test taker is supposedly identified (moderate security).

c. Supervised: human supervision; proctor will login the test taker and confirm correct administration (secure).

d. Managed: high level of supervision with control over the test taking environment through the use of a dedicated testing centre (secure).

In Australia, by 2004 many of the psychological tests used for selection within Defence Force Recruiting (part of ADF) had transitioned to computer-based versions (e.g., the Army General Classification Test – computer version (AGC) (Hinton, 2005)). It was recommended in 2005, however, that the ADF remain in the managed mode of administration for selection tests (by using designated testing facilities) in order to promote reasonable standardisation and eliminate test taker authentication issues (Hinton, 2005).

More recently, Bartram (2010) has proposed a modified model of test administration:

a. Open: unsupervised

b. Controlled: unsupervised

c. i) Remote: supervised

   ii) Local: supervised

d. Fully managed

The application of online monitoring, with real time biometrics, has enabled the emergence of an additional mode of testing (c. i), although this requires the monitoring technology to be available to the test user.

Each mode has advantages. Unsupervised testing is becoming popular as a way for individuals to make decisions about undertaking online therapy programs for anxiety and depression, such as MyCompass (Black Dog Institute) and MoodGym (Australian National University) – see www.mindhealthconnect.org.au for more programs.

Supervision is necessary for high stakes testing such as employment screening. However, techniques have been developed to overcome some of the obvious drawbacks of unsupervised testing. In organisational settings, for example, it is now possible, with some tests, to retest selected (short listed) candidates in a supervised setting using a subset of items from the databank used for the unsupervised testing session, and to compare the results from the two different administrations.

A quick review of publisher test catalogues reveals the impact of computer-based applications for psychological test administration, scoring and reporting. For over a decade, the catalogues from test publishers have reflected the increasing impact of computers and subsequently online testing. Anecdotal evidence from publishers indicates a strong and increasing demand from test takers and test users for tests to be made available online.

The growth of computer-based online testing is discussed further in this document. Advances in technology, and its impact on testing practice, and even test development, indicate the need for ongoing monitoring of developments globally. The widely cited American Psychologist article by Naglieri (2004) provides cautionary comment with regard to the use of online testing, while Hambleton, Bartram, and Oakland (2011) provide a brief overview of the (historical) technical advances, and guidelines and standards for the assessment process. The edited book by Bartram & Hambleton (2006) offers a comprehensive outline of a range of issues, including the perspective of the test taker. However, online testing has expanded significantly since this material was prepared for publication, facilitated by the factors mentioned below.

2. Factors driving the increasing use of online testing

Online testing (a subset of Computer-Based and Internet Delivered Testing) has developed rapidly in recent years, driven by various factors including, but not limited to:

- The rise of globalisation and the increasing need for speed and efficiency in test administration and subsequent decision making.
- Advances in technology, including computer hardware, software and connectivity.
- Increased cost effectiveness and accuracy, through the use of computers and the internet, for both test administration and scoring.
- Cheaper access to the technology, resulting in a significant uptake in computer usage and internet access globally.
- Enhanced capacity for developing a broader range of tests and test items, at times drawing upon advances in modern psychometric testing including item response theory (IRT) and generalisability theory. Such theoretical and computer developments often underpin test adaptation from one culture or language to another.
- Increased opportunity for delivering different item response formats including (dynamic) real time computer adaptive testing, for cognitive, personality and preference tests. This reduces testing time while offering the possibility of enhanced test score reliability and also allows for multiple forms of the same test, reducing practice effects and potential for cheating.
- Enhanced data security (often) and increased speed and efficiency in data transmission and storage.

- The online administration of tests increases the protection of the copyright and intellectual property of the test publishers, thus enhancing publisher acceptance for the online mode of test administration.
- The internet can be used to disseminate material to support test users. This can include online materials such as manuals, FAQs, norms (including updates), practice questions and information for test takers.
- The need to access, sometimes at short notice, test takers in remote locations, often for job selection or high stakes testing purposes.
- Data can be easily and cheaply collected to assist with the development of norms for specific groups or locations.

3. Usage of online testing

3.1 Organisational settings

The 2011 Global Assessment Trends Report (Fallaw & Kantrowitz, 2011) is based upon responses from 463 HR professionals representing companies working with SHL PreVisor. Australasia represented 8% of the sample, and the Americas 39%. Some highlights, bearing in mind the possible limited nature of the sample, are as follows:

- 85% of the companies use testing in addition to other forms of assessment.
- 81% of the companies use online rather than paper and pencil (P&P). (However, the volume of tests administered online is more than 95%.)
- Use of remote (unproctored) testing (commonly referred to as UIT) has increased year on year since 2009. In 2011, 83% of professionals indicated they allowed test takers to complete online assessments remotely. The main reason being convenience for both candidates and test administrators.
- Use of mobile devices for testing is growing and 33% of companies indicated they would allow their use. However, only 10% of companies are requesting that tests are made available this way.

The most recent report (Fallaw, Kantrowitz, and Dawson, 2012) provides similar data. However, Australian psychologists should note that the researchers found regional differences in attitudes towards testing via mobile devices, with job candidates from Asia, as compared to the Americas and Europe/Africa, more likely to request the ability to undertake assessment on mobile devices.

3.2 Educational and other settings

At the 2012 International Test Commission Conference, Martin Roorda (of The Netherlands) delivered a keynote address: *"The Exciting Future of Educational Testing"*. While this is not necessarily the same as psychological testing in educational settings, there is no escaping the overlap between this testing (often achievement testing) and psychological testing in organisational and educational settings. The rise of modern psychometric developments, and enhanced technological applications, may well allow learning diagnostics and processes to be individualised (in what has been termed "The Holy Grail" in education).

From what appears to be a reference to item response theory (versus classical test theory), Roorda referred to "less is more" (i.e., fewer items in a given test for equivalent reliability), real time analysis and evaluation of the educational intervention. Computers, and online testing, are now part of modern educational systems.

As an example of this, Cognitive Load Theory (Sweller, Ayres, & Kalyuga, 2011) posits that instructional materials need to be modified as a learner moves from knowing very little about a topic (novice) towards knowing a lot (expert). Online testing can be used to assess an individual's current level of expertise so an instructor (or computer delivered tutorial program) can decide the optimal design of teaching and learning materials to be subsequently presented to the learner.

It is easy to see similar applications in clinical psychology whereby online test results can be used to provide individualised treatment programs. By making use of item response theory and the power of computers, a branching technique can be employed to provide quick diagnostic outcomes and recommended intervention options for the treating clinical psychologist. Furthermore, with the advent of multi-media simulations, as discussed in Section 9 of this document, it is quite possible that the training and the assessment of provisionally registered clinical psychologists can be facilitated through such online applications.

The use of computerised testing and assessment in education is not new, however. A well regarded book *"Item Response Theory for Psychologists"* (Embretson & Reise, 2000) targets educational and other psychologists. Knauss (2001) commented on computerised psychological testing in her article *"Ethical issues in psychological assessment in a school setting"*. Furthermore, Hambleton (2010) stated that in five to ten years all testing will be conducted online (apart from certain clinical and neuro-psychological applications). Even then, we are seeing online testing applications penetrate areas that, traditionally, were reserved for one-to-one or direct administration of tests used for diagnostic purposes.

This rapid growth of online testing will only be reinforced by developments in China. The huge population, and a lack of traditional testing practice, have driven the uptake of certification testing as well as psychological and educational testing. According to Zhang, Zhang and Zhang (2012), over five hundred academic theses on item response theory have been published since 2001, with computerised adaptive testing (CAT) a "hot spot".

4. Standards, guidelines and good practice

Much of what pertains to good online testing practice mirrors what is regarded as good testing practice in using traditional paper and pencil tests, as outlined in the APS Guidelines for psychological assessment and the use of psychological tests (APS, 2009) and Supplement to guidelines for the use of psychological tests (currently under revision; APS, 1997). In addition, the International Test Commission (ITC) has produced several relevant guidelines designed to promote good practice, with the International guidelines for test use (ITC, 2001) of note.

The following elements are recommended as examples of good testing practice, particularly when the testing is conducted online:

a. Establish which tests are to be used (if any) and the criteria against which test outcomes will be assessed (i.e., is "testing" necessary?).

b. Ensure the test taker is aware of the purpose of the testing and how the test results are to be

used and stored. Inform the test taker of their capacity to receive feedback, and the timing and mechanisms by which this can be achieved. A privacy and consent form is often needed to be signed, particularly in employment and educational setting settings. With children, consent is generally required from both the child and the parent or legal guardian.

c. Clarify the number and type of tests to be administered, and facilitate the opportunity for the test taker to undertake brief practice (sample) items online before taking the test.

d. The test taker should be asked to confirm that they will complete the tests according to the instructions (e.g., not collude with others or seek assistance). Often such an undertaking is required in the introductory phase to the online tests. Research (e.g., Ariely, 2012) suggests reminding people about the need to act honestly diminishes dishonesty. This suggestion aligns with the technique of 'moral suasion', which is used to influence test takers to respond in an honest and transparent fashion.

e. If the testing is to be conducted in an unproctored fashion, encourage the test taker to undertake the tests at a time and location so as to minimise interruptions.

f. Ensure that the test taker has read and understood any email/online instructions for taking the test(s) online. Where the UIT is being used in a medium to high stakes setting for employment purposes, inform the test taker that there is a high likelihood that subsequent confirmatory testing will need to be undertaken under proctored conditions using parallel or similar tests. There is some suggestion also that a test taker knowing of the opportunity to receive personalised feedback may also assist in diminishing malfeasance in UIT.

g. Once the confirmatory testing has been completed, compare the results (automatically calculated and compared by some testing systems) to determine the appropriate course of action (see below: Ethics). (The test user may find value in suggesting to the client organisation that confirmatory testing reflects the existence of high or professional standards on the organisation utilising these tests. This is a positive attribute in itself for many job seekers.)

h. Establish which set(s) of norms is (are) to be used, and whether these are local, global, or both.

i. Store test data and reports in accordance with professional practice guidelines.

An additional key document is the International guidelines on computer-based and internet delivered testing (ITC, 2006). These guidelines provide specific advice for three distinct groups: publishers, developers, and test users, with four general themes addressed, namely:

- Technology – ensuring that the technical aspects of CBT/Internet testing are considered, especially in relation to the hardware and software required to run testing.
- Quality – ensuring and assuring the quality of testing and test materials and ensuring good practice through the testing process.
- Control – controlling the delivery of tests, test taker authentication and prior practice.
- Security – security of the testing materials, privacy, data protection and confidentiality are the four issues and are further broken down into second level specific guidelines, with a third level set of accompanying examples provided to the relevant stakeholder.

5. Advantages of online testing (over traditional or paper and pencil testing)

a. Test Users:

- Developers can embrace the power of modern psychometrics to develop tests which can be adapted cross-culturally (employing techniques such as Differential Item Functioning (DIF)) and which will be more efficient. Ability or cognitive tests in particular can make use of very large item databanks, with items selected randomly for a given level of difficulty. Thus, the early use of computers as merely "page turners" has been supplanted by this method known as linear-on-the-fly testing (LOFT). A more advanced technique involves Computer Adaptive Testing (CAT) where items presented to the test taker vary dynamically according to the correctness of their prior response and until the Standard Error of Measurement (SEM) falls below a pre-defined level (Embretson & Reise, 2000).

- Online tests often provide enhanced security as the problem of inappropriate access to test papers is no longer an issue. (Nevertheless, system access security issues still apply.)

- Publishers can protect copyright and intellectual property as the test items are difficult to copy and the scoring protocols are not revealed. Furthermore, protective item formats (such as the "Foster Item") can be developed so that in a multiple choice test, the test taker has a limited opportunity to be exposed to all response choices for a given item.

- Publishers can take control of a centralised databank, updating norms for convenient distribution to test users.

- Publishers can facilitate the training and education of test users via online mechanisms (including webinars) and take advantage of online enquiries and error messages.

- Malfeasance (or cheating) is an issue for all forms of testing, particularly in testing for high stakes employment purposes. However, online testing can provide the following safeguards (perceived as advantages as well):

  - Keystroke analytics (an example of online biometric authentication)

  - Certified Online Proctoring (e.g., online webcam)

  - Protective item formats

  - Strong machine and browser lockdowns

  - Real time data forensics (e.g., monitoring of response patterns, response latencies, etc which may suggest prior knowledge or attempts to cheat)

  - Unauthorised keystroke monitoring (e.g., issuing of warnings by the proctor for test taker attempts to bypass controls)

  - Following existing security standards, which can include monitoring of web traffic.

- The organisation commissioning the tests is likely to tap into a larger applicant (test taker) pool, and secure a quicker response.
- Practitioners (not all psychologists) have the opportunity to gain quick access to test takers, both locally and remotely. Online testing, whether conducted under proctored or unproctored conditions, does not require the forwarding of test materials by either post or courier, providing a saving of time and expense.
- Publishers can ensure that outdated tests cannot be used as such tests can be withdrawn from the publisher's server.
- Online tests often will be cheaper, faster and better. But not always, and test user skills are still important.
- Scoring is standardised and error free (apart from systematic error in the programming) with data based reports produced quickly. A range of narrative and interpretive reports can also be generated. [However, there are concerns when those untrained in good test usage, and appropriate interpretation, have access to such computer generated reports.]
- Publishers still require test users to meet certain defined qualification levels. While the potential for materials to fall into the wrong hands exists, this problem is unlikely to be any more widespread than is the case with paper delivery.

b. Test Takers

Increasingly test takers appear to appreciate having the opportunity to undertake tests in a familiar, home environment, using technology and equipment with which they feel comfortable. It is convenient, particularly for those who are not working in an urban or major regional centre, or those who find it difficult to undertake testing during normal business hours. UIT is used extensively in the resource sector, where test takers may be working remotely and/or operating on a Fly-In-Fly-Out (FIFO) basis.

6. Issues and potential disadvantages of online testing

Research into the efficacy of online testing and the balance of risks and rewards is relatively new. However, the following elements have been raised by researchers:

- A paper and pencil test, converted to an online format may possess different psychometric properties from that of the original test. Both construct and measurement equivalence are required. Appropriate piloting and/or simulation needs to be conducted, with a focus on matters such as Differential Item Functioning.
- Research using personality questionnaires suggests that there is very little difference in outcomes between UIT administered tests versus proctored internet tests, even for high stakes testing (for example, Bartram & Brown, 2004, in their research using the OPQ). However, Guion (2011) has expressed doubts and wonders if their results are typical. Moreover, issues can exist for online ability tests conducted for medium to high stakes purposes. A key area of focus is in relation to the test taker, including not only authentication and cheating concerns but also how UIT may affect individual test takers and their attitudes towards a potential employer.

- Cheating on cognitive tests (as opposed to faking or response distortion on non-cognitive measures) can be an issue for UIT. While "speeded" high stakes cognitive tests appear to be partially buffered from the cheating phenomenon, "power" tests are likely to be more vulnerable. Macqueen (2012) cites two presentations from the 2012 SIOP Conference in which the estimated base rate of cheating is claimed to be low. However, what level of confidence is required for one to conclude that a test taker has cheated when a verification score differs statistically from the original UIT score?

- Surrogates may undertake the tests, although authentication can also be an issue for traditional testing. Another scenario, difficult to monitor, is when an accomplice is positioned near the test taker, but beyond the view of a webcam, even if one is being used.

- There has been some support for the view that older test takers, unfamiliar with computers and technology, are disadvantaged by the use of timed tests in high stakes testing by UIT. No gender differences appear to operate, although there appear, in one recent study at least, to be demographic differences in the test takers' perception of the testing environment. Furthermore, the environmental trade-off between proctored onsite and unproctored administration appears to be better workspace versus less noise, respectively.

- Despite the above, UIT is likely to be associated with greater variance in the testing environment. Under traditional, proctored testing practice, a test administrator can control many external factors and/or make note of any anomalies that may have affected the test taker's performance or responses. The increasing use of internet cafés or the use of internet connections in airport lounges is not conducive to delivering an optimum performance for the test taker. The advent of test delivery on mobile devices increases the likelihood of variability in the testing environment. In addition, poor internet connectivity can have an adverse effect on the testing environment.

- Online testing is often accompanied by a complete lack of interaction between the test taker and the psychologist (or professional test user). This may compromise the quality and comprehensiveness of the assessment judgments and subsequent decisions. Important non test personal information may be overlooked, as may relevant contextual factors.

(For further information see Tippins (2009), together with subsequent commentaries; and Bartram (2008).)

7. Technical issues

It is important for all groups of test takers to have equality of access. This not only has implications for the test design and content, but also for the technology used to deliver the test. The ITC (2006) Guidelines (Guideline 1) provide the following assistance:

"Give due regard to technological issues in Computer-Based (CBT) and Internet Testing:

a. Give consideration to hardware and software requirements.

b. Take account of the robustness of the CBT/Internet test.

c. Consider human factor issues in the presentation of material via the computer on the internet.

d. *Consider reasonable adjustments to the technical features of the test for candidates with disabilities.*

e. *Provide help, information, and practice items with the CBT/Internet test.*

The advent of updated internet browsers and the presence of applications designed to protect the computer can sometimes mean that the testing system fails to load or run appropriately. Variations in internet connection speed, the operating system and the browser need to be considered at the development stage. Furthermore, "maintenance" issues are particularly important for test publishers. The ITC (2006) Guidelines provide specific guidance, but some publisher systems or platforms appear to be more user friendly than others. The more problematic systems build in a great deal of redundant protection, with a complex randomly generated password (and a suitable but not necessarily obvious ID). Such passwords can be transmitted and/or entered incorrectly if the test administrator or the test taker is not careful, leading to subsequent test taker frustration with the testing process.

8. Ethics

Apart from standard ethical practice as it applies to any testing or assessment, online testing, particularly UIT for high stakes testing, brings to the frame the key issue of malfeasance or "cheating" and what to do about it if it is detected or suspected.

The existence of cheating is likely to lead to inappropriate (job) selection decisions being made when UIT testing is used in high stakes situations. Thus, there is a need to confirm the results through some process such as a subsequent proctored administration of a parallel form, or similar test. However, there is a clear ethical and professional issue involved here: At what level of discrepancy (between the two test scores) can the test user claim conclusively that cheating has taken place? What confirmatory evidence is available to support the conclusion and what does the organisation (or hiring manager) do about it? Is procedural justice ignored if the person has no counter-claim available? What are the risks involved for the major stakeholders, and how should these be managed?

To reduce the probability of being caught in this dilemma, prevention is important as has been noted in previous sections on test security, as well as the need to inform the test taker of the procedures. Some organisations may even employ an explicit honesty policy before testing commences.

When a given number of people are to be employed through a large scale testing and selection assignment, a cut score approach may be employed. However, instead of using a simple top-down selection approach, it is recommended that the test user initially selects more test takers than anticipated for the second, confirmatory, testing phase. To the extent that cheating occurs, the number passing the cut will be higher than expected, but the additional numbers will be eliminated by the confirmation test. It should be noted, however, that in Australia a great deal of testing involves smaller groups, including individual assessment. However, graduate recruitment programs, and other large scale selection programs, should consider employing this modified cut score approach in order to reduce the impact of cheating (Bartram, 2009).

Even if currently the extent of cheating in UIT is relatively small (as suggested by the research of Guo, Drasgow, and Gibby (2012) and Weiner and Rice (2012)), good practice demands that some form of proctored testing is conducted before a final decision (or diagnosis) is made, particularly in high stakes testing.

9. Future developments

Online testing is expanding rapidly, particularly with the convergence of technology and the acceptance of "connectivity" as part of life for the vast majority of adolescents and adults within our society. The rapid growth in information exchange via digital means will further the drive towards online testing and assessment. Test takers can now complete personality questionnaires via mobile devices. It is understood that test publishers are responding to the demands of consumers (test takers) in developing such applications for mobile devices.

Furthermore, apart from ease of use, technology provides the opportunity to develop and present richer forms of stimuli than is possible with paper and pencil or traditional testing. Such developments can incorporate audio, video and graphical stimuli. Greater realism can thus be provided than is possible through a written scenario. Technology can provide more standardisation than is possible with live role plays (even if professional actors), a traditional practice or activity in comprehensive assessment centres used for selection and development purposes. (A description of video-based testing at US Customs and Border Protection is provided by Cucina, Busciglio, Thomas, Callen, Walker, & Goldenberg Schoepfer (2011)). Use of such technology-enhanced testing is not restricted to management levels, with examples existing for the use of technology to assist in the assessment of unskilled or semi-skilled personnel, particularly those challenged with literacy issues. Such developments can combine animation with graphical tools such as drag-and-drop controls (Reynolds & Dickter, 2010).

The term "gamification" has entered the testing and assessment lexicon. Software applications may include animated avatars and simulated environments. While downloadable games such as America's Army probably have more to do with recruitment and public relations rather than testing per se, the concept is gaining increasing traction, including within the educational sphere for learning and assessment purposes. The opening state-of-the-art speech at the ITC 2012 conference was titled "The evolution of assessment: Simulations and serious games" (Fetzer, 2012).

The above suggests that there is an increasing blurring of the lines between "tests and testing" and other forms of "assessment". There is a range of issues to address, regardless of the popularity in adopting such technological innovations. "Construct equivalence" is a particularly important technical issue to address as are professional issues such as the confidentiality and security of information. Furthermore, what opportunities are provided for proper test taker feedback when automation is the focal point? In addition, automation can mean that the test taker's micro behaviours can be recorded during a computer-delivered assessment. Metrics such as click patterns or mouse "hover-time" may be collected, with the possibility of reductionist or spurious assessment judgments being made without the support of adequate research (Reynolds & Dickter, 2010).

10. Implications for the education, training and professional development of psychologists in Australia

The current Australian Psychology Accreditation Council (APAC) educational requirements for testing and assessment competence provide limited guidance in the area of technology and psychological testing; and the psychometrics underpinning modern test developments. (Note, however, that these guidelines are in the process of being reviewed at the time of preparing this document.) Similarly, CPD and related initiatives in Australia appear to offer very little for practitioners wishing to develop their testing and assessment skills. Publishers can provide limited training (relevant to the operational elements of a given test or testing platform), but the broader underlying principles and issues are not canvassed in depth.

The lack of focus in this online area (of testing and assessment) in Australia appears to be associated with a lack of research in the testing and assessment domain, as well as a lack of CPD, even at major conferences. For example, at the 2011 biennial APS IOP Conference, with 600 registrants, there were no presentations on technology and testing and perhaps only one or two in the testing domain as a whole. This contrasts with what is happening overseas, where the annual SIOP conference (4,500 registrants) has a solid focus on testing, associated technology developments, and the implications for psychologists (test users) and test takers.

Furthermore, the 2012 ITC Conference had as its theme: "Modern advancements in assessment: testing and digital technology, policies, and guidelines". This theme was to reflect the changes that have occurred over ten years since the 2002 conference in Winchester (UK), with its theme: "Computer-based testing and the Internet".

At this stage, the APS Tests and Testing Reference Group (TTRG) has been established to address "Tests and Testing". However, it should be noted that technology is blurring the lines between testing and other forms of assessment. Perhaps in recognition of this, the European Federation of Psychologists' Associations (EFPA) restructured in late 2011. As a result, instead of having a "Standing Committee on Tests and Testing", the EFPA now has a "Board of Assessment".

11. Conclusion

While online testing, at this stage, is most relevant to organisational and educational psychologists, it does impact on many potential test takers (and organisations) in Australia. In addition, advances in "technology-enhanced" assessment will also need to be monitored and addressed. This is apart from the recent release of ISO 10667 relating to workplace assessments. This ISO standard addresses all forms of work related assessment, including psychological testing. (The implications of this ISO standard are yet to be determined. As of July 2012, there appear to be no implementations of this standard in any countries, including parts of Europe where ISO 10667 has been supported strongly.)

Online psychological testing is here to stay, and that includes UIT. Psychological testing via online devices is widely accepted (and even expected) by the broader community and this is evidenced by the statistics revealed by one test publisher/consultancy at SIOP 2012. Of 8,000 candidates tested per day, 65% were tested under UIT conditions. (The extent of follow up verification is unknown.) At the same conference it was reported that a major US agency, the Office of Personnel Management, has been instructed to introduce UIT.

Psychological testing has historically been viewed, in the main, as being the province of psychologists. While this claim may be debated (for example, by some educationalists), technology has been a significant catalyst in changing the dynamics and speed of the testing process over the past decade. Given the significant global uptake of (and demand for) online psychological testing, it will be important for Australian psychologists to gain advanced psychometric, testing and assessment skills while simultaneously being effective in educating their client base regarding the benefits and limitations of online testing. In essence, psychologists will need to demonstrate their capacity to "value add" well beyond what is offered by cost effective and streamlined online testing systems.

12. References

Ariely, D., (2012). *The (Honest) Truth about Dishonesty: How we lie to everyone – especially ourselves.* New York: Harper.

Australian Psychological Society (2009). *Guidelines for psychological assessment and the use of psychological tests*. Melbourne: Author.

Australian Psychological Society (1997). *Supplement to Guidelines for the use of psychological tests*. Melbourne: Author.

Bartram, D. (2010). *The need for a new mode of test administration in assessment for work.* Paper presented at the 7th International Test Commission Conference, Hong Kong.

Bartram, D. (2009). The International Test Commission Guidelines on Computer-Based and Internet-Delivered Testing. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2(1)*, 11-13.

Bartram, D. (2008). The advantages and disadvantages of on-line testing. In Cartwright, S. & Cooper, C. (Eds.), *The Oxford handbook of personnel psychology.* Oxford: Oxford University Press.

Bartram, D. (2001). *The impact of the Internet on testing for recruitment, selection and development*. Keynote paper presented at the Fourth Australian Industrial and Organizational Psychology Conference, Sydney.

Bartram, D. (1987). The development of an automated pilot testing system for pilot selection: The MICROPAT project. *Applied Psychology: an International Review, 36,* 279-298.

Bartram, D., & Brown, A. (2004). Online testing: Mode of administration and the stability of OPQ32i scores. *International Journal of Selection and Assessment, 12(3*), 278-284.

Bartram, D., & Hambleton, R. K. (Eds) (2006). *Computer-based testing and the internet: Issues and advances*. Chichester: John Wiley & Sons.

Cucina, J. M., Busciglio, H. H., Thomas, P. H., Callen, N. F., Walker, D. D., & Goldenberg Schoepfer, R. J. (2011). Video-based testing at U.S. Customs and Border Protection. In: Tippins, N. T., & Adler, S. (Eds). *Technology-Enhanced Assessment of Talent*. San Francisco: Jossey-Bass.

Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Fallaw, S. S., & Kantrowitz, T. M. (2011). *2011 Global Assessment Trends Report*. SHLPreVisor.

Fallaw, S. S., Kantrowitz, T. M., & Dawson, C. R. (2012). *2012 Global Assessment Trends Report*. SHL.

Fetzer, M. (2012). *The evolution of assessment: Simulations and serious games.* Paper presented at the 8th Conference of the International Test Commission, Amsterdam.

Guion, R. (2011). *Assessment, Measurement, and Prediction for Personnel Decisions* (2nd Edition). New York: Routledge.

Guo, J., Drasgow, F. & Gibby, R. E. (2012). *Estimating the base rate of cheating for unproctored Internet tests.* Paper presented at the 27th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego.

Hambleton, R. K., Bartram, D., & Oakland, T. (2011). Technical advances and guidelines for improving test practices. In: Martin, P. R., Cheung, F. M., Knowles, M. C., Kyrios, M., Littlefield, L., Overmier, J. B. & Prieto, J. M. (Eds). *IAAP Handbook of Applied Psychology.* Chichester: Blackwell Publishing Ltd.

Hinton, M. (2005). *Review of the use of Internet-based testing in recruitment and selection.* Department of Defence. Canberra: Australian Government.

International Test Commission (2001). International guidelines for test use. *International Journal of Testing, 1(2),* 93-114. Retrieved 22 January, 2013, from http://www.intestcom.org/upload/sitefiles/41.pdf

International Test Commission (2006). International guidelines on computer-based and internet-delivered testing. *International Journal of Testing, 6(2)*, 143-172. Retrieved 22 January, 2013, from http://www.intestcom.org/Downloads/ITC%20Guidelines%20on%20Computer%20-%20version%202005%20approved.pdf

Knauss, L.K. (2001). Ethical issues in psychological assessment in school settings. *Journal of Personality Assessment, 77(2)*, 231-241.

Macqueen, P. S. (2012). The rapid rise of online psychological testing in selection. *InPsych, 34*(5), 16-17.

Moreland, K.L. (1992). Computer-assisted psychological assessment. In M. Zeidner & R. Most (Eds.), *Psychological testing: An inside view*. Palo Alto, CA: Consulting Psychologists Press.

Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A. L., Margolis, A., & Velasquez, R. (2004). Psychological testing on the internet: New problems, old issues. *American Psychologist, 59(3)*, 150-162.

Reynolds, D. H., & Dickter, D. N. (2010). Technology and Employee Selection. In: Farr, J. L. & Tippins, N. T. (Eds). *Handbook of Employee Selection.* New York: Routledge.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive Load Theory. New York: Springer.

Tippins, N. T. (2009). Internet alternatives to traditional proctored testing: Where are we now? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2(1)*, 2-10.

Weiner, J. A., & Rice, C. (2012). *Utility of alternative UIT verification models.* Presentation at the 27th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego.

Zhang, J., Zhang, M., & Zhang, W. (2012). Application and Development of Testing in China. *Testing International, Vol. 27, July, 2012 (6 – 8).* [Newsletter of the International Test Commission.]

For more information about the APS please visit **psychology.org.au** or contact:

The Australian Psychological Society Limited
PO Box 38, Flinders Lane, VIC, 8009

Telephone: (03) 8662 3300 or 1800 333 497
Fax: (03) 9663 6177
Email: contactus@psychology.org.au

ABN 23 000 543 788

18APS-PP-B-OPT-P1