Using online assessment tools for recruitment

PTC

Psychological Testing Centre
www.psychtesting.org.uk

Contents

Introduction

The growth in the internet over the last decade has been phenomenal. In the UK it has moved from being the domain of academics and IT professionals to a medium accessed by over 71 per cent of the working age population (based on Office of National Statistics findings). About a quarter of internet users cited looking for work as one of the activities they performed on the web. The growth of internet recruitment sites supports this trend. Most large companies have recruitment pages on their website and many allow applicants to complete applications forms or submit CVs. A growing number also use the medium for more structured assessment processes. It is now not uncommon for candidates to be asked to complete tests and questionnaires online, particularly for initial sifting purposes. This guidance looks at some of the good practice and ethical issues in this development in the use of tests.

As with other uses, while the internet is very useful in easing communication and managing data, it does not generally change the essential nature of a transaction. Buying goods over the internet is not very different from old-fashioned mail order. Submitting an application via the internet is not very different from sending it in the post. However, use of the internet has encouraged a form of remote testing which was perhaps less common before the internet era. Biodata forms and other questionnaires were sometimes sent out by post to be completed together with application forms for initial sifting purposes. Automated telephone systems were occasionally used to collect applicant data. Now it is quite common for a candidate to complete a personality questionnaire online early on in the selection process and the use of remote online ability measures is also growing. Thus when considering the impact of the internet on testing practice the issues can be divided into those related to remote assessment, those related to computer administration and those related to early stage sifting and shortlisting. These are all issues that existed before the internet but have become more pertinent with its appearance. The APA committee on internet testing came to a similar conclusion. Their report (APA, 2004) uses the phrase 'new problems, old issues' in its title.

Remote assessment

The ITC Guidelines on Computer-Based and Internet Delivered Testing (2005, copies available from the PTC website, www.psychtesting.org.uk) defines four modes of test administration: Open, Controlled, Supervised and Managed. Remote testing in a sifting scenario is most likely to be in 'Controlled' mode – that is there is no direct supervision of the testing session but test-takers will need to register as applicants and perhaps wait to be sent a URL and password to enable them to log on to the testing site. Less frequently 'Supervised Mode' may be used – candidates take the test at a supervised site where an administrator can verify their identity, log them on and control testing conditions, but leaves the administration to the internet. 'Managed' mode is similar to 'Supervised' mode but involves a higher degree of control over the test-taking conditions. 'Open' mode, where the candidate does not even necessarily have to register, is not useful for sifting since the candidate cannot be identified, but it is used by some employers for assessments aimed to promote self-selection. Candidates may be invited to take a test and then provided with feedback regarding their suitability. Good candidates would be encouraged to apply and those not meeting the criteria informed of their low probability of success. However, these results would not be used by the employer in making selection decisions at any stage.

Remote testing was possible before the internet. However, this required test users to send out a test to the applicant, typically by post. The candidate would then fill it in and send it back to be scored. A key difference between this and the internet, is that for internet-based remote testing, there is more control over both the materials and at least some features of the administration. When materials are sent out by post, all control by the user over what subsequently happens to them is lost. With the internet, the system implementation can manage and monitor the interaction with the test taker. When test users purchase access from a test distributor they need to make sure that the degree of control and candidate management suits their needs.

Whether or not the internet is used, remote administration raises a number of good practice issues.

Test security
Without direct supervision, it is easier for test content to be copied, widely distributed and, particularly in the case of ability tests, compromised. With internet test administration it is possible to take some technological measures to impede the ease with which test content can be copied or downloaded. For any test administered in 'Open' or 'Controlled' mode the user needs to take into account that the content could become widely distributed and monitor as far as

possible whether this is happening (e.g. by checking appropriate internet chat groups and websites.) In 'Controlled' mode there is an audit trail that can be followed if materials are compromised. However, this is not possible for tests made available in 'Open' modes. For ability tests (and other instruments with 'right' answers), the availability of multiple alternate forms which can be used if a test becomes compromised should be considered as essential for any test that is to be used unsupervised in a high stakes setting.

For ability testing, an approach to this problem is through the use of test generation techniques to create multiple equivalent test forms. If each test taker receives a different set of test items and if the pool of items from which the tests are constructed is regularly updated, then there is little value in trying to cheat by 'harvesting' the test items.

Informed consent

It is important that anyone undertaking a psychological test understands what it entails and why it is being used. Under the Data Protection Act, information should only be collected and used appropriately for the agreed purpose. With remote test administration, the process should include a clear explanation of the nature of the test and how it will be used in the selection process so that the candidate understands what will be revealed by the test and can enter into the process with a proper understanding. Typically in a selection situation the candidate is deemed to have provided consent to use of the information by the act of providing it. However, with a psychometric tool, it may not be transparent to the candidate what is being assessed.

Adapting for disability or language

Candidates with disabilities or who are not primary speakers of the language of the test may not be able to respond to the test or may not do themselves justice in responding with a standard administration. It is important that the introductory information allows such candidates to identify that there may be a problem and assess whether taking the test in the standard way will be appropriate for them. For disabled candidates for whom travel is difficult, remote testing that allows the completion of the early stages of selection at home can be a boon.

A good test will be designed with accessibility in mind – see 'computer administration' section – however, even with the best design, there are likely to be some candidates for whom the test is not sufficiently accessible. There should be a clear process for candidates to follow if they wish to check if the test is appropriate for them as well as an alternative selection track for those for whom it is not. It would be contrary to the Disability Discrimination Act (1995) to exclude candidates with disabilities in this way.

Candidates should have access to information about the nature of the assessment and access to sample or practice materials. These should be as like the actual test as possible so candidates can judge if they are suitable (e.g. if the font size is suitable or whether they can see figures clearly). There should be an explanation of what to do if candidates feel that they might have difficulty with completing the test because of a disability, or because of poor language skills. A specific point of contact is preferable to a vague recommendation to 'contact us'. Where candidates are invited to take the test via e-mail or letter this should refer candidates to practice materials and encourage them to ask for accommodations if they are needed. This is particularly important because there will be no test administrator to identify that a candidate is having difficulties and prevent an inappropriate test administration.

One of the benefits of recruitment via the internet is its wide reach. However, when it comes to assessment, this may mean that an instrument, which is designed, and trialled for local use, is accessed by more distant applicants from different linguistic or cultural backgrounds. Users need to consider their policy with regard to applicants from unexpected places and whether the test is appropriate. The linguistic diversity of job applicants today means that any selection process needs to consider how assessments will work with those whose primary language is not English. When the language demands of the assessment exceed those of the job, there may be a problem. This could happen, for instance, with a written test for a job, which only required spoken English skills. One of the benefits of the internet is that where a test exists in multiple languages it is relatively easy to allow candidates to select to do the test in the most appropriate language for them.

Test administration
The importance of good test administration procedures is paramount to test effectiveness. Lack of standardisation can reduce the reliability and validity of test results. Timed ability tests are particularly sensitive to variation in administration conditions and procedures. While online delivery increases the degree of precision with which timing and other factors can be managed, remote administration reduces the level of control that can be exercised over the test-taking environment. If not properly briefed, the candidate may take the test in an environment where he or she may be distracted by external noise or interrupted by phone calls while taking the test. While such problems can occasionally occur in supervised modes, they are less likely in a well-managed testing environment and would be dealt with by the test administrator.

Candidate instructions should explain the importance of taking the test in the right mental state and in the right environment, trying to avoid interruptions and distractions, e.g. by turning off mobile phones, not working in a busy noisy

environment and so on. Encouraging candidates to make use of practice or sample materials will help them to understand the nature of the assessment and appreciate the importance of completing the test in appropriate surroundings. As with any other assessment situation information also helps alleviate anxiety and promotes impressions of procedural fairness.

Another function of test administrators is to check that candidates have understood the test instructions correctly and that testing is proceeding appropriately. In the remote situation, candidates who fail to understand the test requirements, whether because they skip the instructions or because their language skills are insufficient may not do themselves justice on the test. The test design should include checks that candidates understand what they have to do. This might be through example questions that candidates have to answer correctly to proceed, or interactive instructions, which require candidates to read and understand the text. On-screen help should be accessible during the test as well as before it starts.

Control of fraud

In considering the issue of deception and cheating, it is important to distinguish between tests of maximum performance (e.g. ability tests), which have right answers, and tests of typical performance (e.g. personality questionnaires), which do not.

For tests of maximum performance candidates may be able to cheat in any number of ways – looking up information, altering the timing mechanism, obtaining the questions in advance. While cheating can occur in supervised testing settings, without a test administrator acting as an invigilator, candidates are much freer to try to subvert the assessment process. Testing systems can be set up to deter if not completely eliminate these various types of cheating. Test generation has already been mentioned as a means of reducing the risk of the answers becoming known in advance. In addition software can include tamper-detection and other tools to indicate when people are behaving in an aberrant manner.

The easiest way to cheat is probably to seek assistance from someone else or to have someone else to take a test on the candidate's behalf. Even when iris-recognition or other similar technology to verify identity is readily available, it will require CCTV monitoring (e.g. using a web-cam) to detect the presence of an additional person providing the answers for the candidate to enter. While this is technically feasible, and is provided as a high security option by some assessment providers, test users have to consider the cost-benefit equation. Currently fully monitored remote testing is expensive (though probably not as expensive as traditional supervised testing). The added cost may be seen as being of little added value unless cheating was a serious problem.

With tests of typical performance, such as personality inventories or integrity tests, the issues are a little different. While people may try to 'fake good', this is something one can do as easily under supervision as on one's own. Remote testing perhaps offers the candidate longer to think about responses, but the issue of distortion of responses will be similar with supervised or unsupervised use of these instruments. The collusion which remote testing allows is less likely to be helpful for tests of typical performance and may well alter results in systematic ways which could be flagged by distortion scales. Slow response latencies can be picked up in computer administration and may also be indicators of insincere responses.

Placing a remote assessment in an appropriate context for the candidate can help deter distortion and cheating. The mutual benefit of accurate testing should be explained to the candidate before embarking on the test. A correct result will help the employer identify the best candidates, but also help candidates avoid positions to which they would not be suited. Candidates should be made aware that those short-listed may be re-assessed later in the selection process under supervision to verify their results. This will lessen the utility to the candidate of gaining a good score by deception. Long-term use of remote tests without later verification of results is likely to undermine the credibility of the instrument and encourage 'cheating support' systems to develop. Verification of results can range from administration of parallel tests to confirm earlier test scores to structured interviews designed to check on the validity of personality profile or other bio-data.

Candidate support

A well-designed internet administration will provide clear information in an easy to follow format and anticipate most candidate questions. However, when things go wrong or candidates have special needs, more interactive support is needed. A helpdesk accessible by phone or e-mail can deal with questions and its very existence may provide reassurance for candidates who are unfamiliar with the medium or who are lacking in confidence. There should be clear information about opening hours and response times. For employers who are concerned about attracting and keeping good applicants, or maintaining a positive image with candidates who may also be clients, good quality support is essential.

Data security

In remote testing, information is being passed between the candidate and the recruiter. The Data Protection Act (1998) requires the recruiter to maintain good security on candidate information whether it is kept in paper or electronic form. Many people are more concerned about data security when it is passed via the internet than when it is in paper form. In either case the system should be designed to maintain high levels of security. Locked storage is required for paper records. With the internet equivalent security includes using secure servers, limiting access to data to appropriate personnel, password protection of data and encryption as appropriate. Data should also be deleted once it is no longer needed, or anonymised if being saved for research purposes.

Computer administration

Equivalence

It is still the case that the majority of tests administered on computer, whether via the internet or not, are adaptations of earlier paper and pencil instruments. Only a minority take advantage of the possibilities that computers offer for interactive displays, complex statistical analysis of responses such as use of item response theory, or automated construction of test forms or items. Test equivalence is, therefore, an important consideration. Timed ability and skill tests can be particularly sensitive to variation in the mode of presentation (e.g. in relation to timing, screen design and ergonomics) especially those that are highly speeded or that rely on complex information displays. If there is clear evidence that computer-administered and paper versions of tests are equivalent in terms of validity, reliability, scale means and standard deviations, then test norms for the computer administered version can be based on existing data (e.g. from paper and pencil administrations).

It should also not be assumed that findings from supervised administration will generalise to remote administrations. It is important to ensure that equivalence has been established in terms of mode of administration and not just in terms of medium. Thus, it may be that a test is equivalent when the paper version (supervised administration) is compared with a supervised computer-based version, but not if the computerised version is administered remotely in controlled mode.

The majority of equivalence studies to date show that psychometric properties are maintained when tests are transferred to computer and for remote administration. However, where there is a strong element of speed, or where item display or response modalities have changed between paper and computer forms, differences are found.

Test consistency

A particular issue with computer-based tests is maintaining the consistency of delivery and display. The hardware and the particular software settings being used can affect speed of delivery, screen resolution, colour values and clarity of display. This means that it is difficult to control what the candidate will see. Some may see a whole item on screen, others may need to scroll to find the answer options or a 'next' button, graphics may be distorted and unclear, and colour contrasts may be affected. Browser type and settings also impact these factors and can even change the font and background colours seen. Other settings and software affect whether particular applications can be used for display (e.g. Flash, Java). The design of the testing system needs to take these factors into account either by controlling them (e.g. by accessing computer

settings and downloading an appropriate test version) or testing that a single version will work equivalently with all settings. This will be easier for a text-based questionnaire than for an exercise, which is based on graphics. In practical terms this means that most self-report inventories are relatively immune to problems of download time, screen layout, browser version and so on. However, ability tests that rely on graphical item displays and make assumptions about what would be visible without scrolling, are more of an issue.

For such instruments, software should check the recipient's computer for compatibility, and the test should only run if the right configuration of hardware and software is found.

A further difficulty in this area is the constant development in technology. A system that has worked well to date may not do so when a new operating system is introduced or typical computer specification is enhanced. Therefore, the user must ensure that system maintenance includes timely updating for new technology. Of course many candidates will not be working from the latest technology, so the system also needs to be backward compliant and equally accessible to someone using a 56K modem on a dial-up line as a fast connection on a ram-rich computer. The supplier needs to check that the system works effectively for the range of software and hardware likely to be used by candidates and test users need to make sure they are aware of any limitations or restrictions.

Technological support

In order to run an effective internet testing process it needs to be based on robust software running on adequate hardware. The test should download effectively and consistently on demand. The server needs to be able to cope with maximum demand levels. In recruitment contexts where there is a closing date, a large proportion of applications are likely to arrive in the last few days and even hours. Where problems occur these need to be addressed quickly. A certain degree of redundancy in the system is desirable. For instance, if a system has duplicate servers, when one goes down or needs maintenance, the other can continue service to candidates. Users should ascertain that the capacity of the system meets their needs and that the supplier can maintain high levels of availability for the system. The user should also check what level of support can be expected, for instance, whether there will be access to a 24/7 helpdesk for information and to discuss problems with the technology.

Contingency design

Even with the best systems problems can occur. It is important to consider what message the candidate might receive in various situations. If the server is busy or another problem occurs the candidate should receive a polite, meaningful and helpful message – rather than a failed connection or a hung up screen. If there is a problem at the candidate end – e.g. the internet connection goes down – there should be a facility for the candidate to reconnect. The candidate will need to be informed what to do next and whether previous responses have been correctly received or lost in the ether. A candidate helpdesk needs to be able to answer these kinds of queries as well as questions about the content of the test or its place in the selection process.

Accessibility

While it may not be possible to make a test accessible to all candidates, good design can make tests more accessible and bad design can restrict accessibility. The Web Accessibility Initiative of the W3C consortium provides guidelines and simple checklists on its website (www.w3.org/WAI) to help designers make applications more accessible. Other organisations also provide helpful guidelines. Simple tips like using the 'alt' text function to describe graphics and choosing appropriate colour schemes and fonts can make the application more accessible and often more user-friendly for everyone. Users should check the accessibility of the application and ensure that candidates can easily check whether they will be able to access the test and have information about what to do if they cannot or are not certain.

Internet access

It should not be assumed, even in these days, that everyone has access to the internet. Where recruitment is only via the internet, certain groups may be excluded. Surveys suggest that people in lower socio-economic groups, older people and people in some regions of the country are less likely to have access to the internet at home and less likely to have used it at all. Employers need to take this into account and in many cases it would be advisable to provide alternate channels for applications apart from the internet.

Sifting and Shortlisting

The third strand of issues is around best practice in sifting and shortlisting with tests and questionnaires. These topics should be covered in test training and are dealt with in general guidelines on testing such as the ITC International Guidelines for Test Use and the AERA/APA/NCME joint standards for educational and psychological testing.

Validity

The most important factor in effective use of a test in any situation is its validity for the intended purpose. In the case of sifting this means that it should measure a job relevant characteristic or set of characteristics and this should be determined through job analysis and matching to test content. Where possible evidence of relevant criterion-related validity should also be sought and follow-up studies organised to check on the test in use. Ideally this should verify that there is a clear relationship between test scores and later job performance. Criterion related validation should be able to show that success rates for those in rejected score ranges are lower than for those in passing score ranges.

However, a sifting instrument can be used for two reasons. Firstly it can be used to predict later job success directly. Alternatively, and more typically, it can be used to predict decisions later in the selection process. A good sifting instrument can increase the proportion of successful candidates among those subjected to more detailed assessment and hence reduce the number of candidates that need to be seen at later selection stages. Here, the main value of the sift is on saving cost while maintaining the overall quality of selected candidates rather than on increasing the quality of those finally selected. Internet based assessments are particularly common in large scale recruitment processes where the cost-savings can be very significant and the potential for statistical validation is high. In this situation, validation is also particularly cost effective because it can help improve the way the test is used (e.g. by optimising score weightings). Even minor increments in validity can have substantial utility with large numbers of positions to be filled.

Use of cut-scores

The nature of sifting requires cut-scores to separate those rejected from those chosen to proceed into the later stages of selection. In managing numbers, recruiters have two options. On the one hand they can select a fixed number of people for the second stage of selection (i.e. use a top-down process of selection). On the other hand they can select all those who pass a pre-defined cut-score. In general, high cut-scores and use of a single criterion or test in shortlisting should be avoided unless there is very strong evidence to support them. A key point in sifting is that it is a process of ruling out people who have a very low probability of success in a job rather than selecting in those with high one.

In more formal terms, we want to minimise the degree to which a sift rejects people who would be suitable for the job while maximising the degree to which it rejects people who would not. The optimal place to set a cut-score to achieve this depends on the base rate (i.e. the proportion of potentially successful candidates in the applicant pool) and the validity of the sift tool. As a rule of thumb it is probably safe to sift out up to 30 per cent of people using a sift tool with good validity. Higher cut-scores increase the risks of rejecting large numbers of potentially good applicants and also increase risks of adverse impact arising if the sift is based on attributes that have different mean scores for different applicant groups.

So, while it may be convenient, particularly when there are many applicants and few positions, to use an online test to screen out a large proportion of candidates, it is unlikely that this practice can be justified. A single measure is unlikely to carry enough information to reject more than a third of applicants, particularly when it is not possible to control the conditions under which the test was completed. Where a greater proportion needs to be screened out, the recruiter should consider a multi-dimensional approach. A questionnaire that addresses a number of job relevant traits or a combination of different measures is to be preferred. Tests can also be used in conjunction with other gross negative disqualifiers, such as the absence of a necessary qualification, to broaden the sifting criteria. The multidimensional approach takes into account several different aspects of performance. A proportion of candidates will be very weak in every area and, taking all the measures together, a larger proportion of candidates can be screened out as being unsuitable without using high cut-offs on any one instrument. This approach can also increase the validity of a selection procedure by assessing more aspects of performance potential.

Reporting and feedback

Online systems are often supported by automated reporting facilities, particularly where decisions are not based on a simple cut-score rule. In evaluating a system, due regard needs to be given to the quality and accuracy of the output, particularly where there is a substantial interpretive element. Poor reports can undermine the value of a good instrument just as a report that seems well written may be invalid if based on the results of a poor measure.

Reports are typically used by recruitment consultants or hiring managers who are not trained test users. As lay users they need clear, non-technical information in a form that directly addresses the issues they are concerned with. In recruitment, hiring managers are concerned with the risks associated with a candidate in terms of likely fit or lack of fit to the job requirements. Reports should make clear the status of the information provided, in terms of the confidence that can be placed in it, and should always stress the value of corroboration through the use of multiple sources or methods of assessment.

Similar considerations should be taken into account when designing reports for providing feedback to candidates. Candidates should always be provided with access to a qualified test user if they have any concerns or issues they need to raise that lie outside the competence of the hiring manager.

Test quality

The psychometric properties of a test or questionnaire need to be of a proper standard for the use in hand. This means that items should be well constructed and the test should have an appropriate level of reliability in the applicant population for the suggested use. There should be suitable norms available. It is the nature of the internet that many systems are developed by people with IT and marketing skills rather than expertise in assessment. While a test may look good on screen, generate attractive output and even have some face validity, it may not be a technically sound instrument. If it is not supported by appropriate technical data, the user cannot assume a test has good measurement properties.

The British Psychological Society has introduced a Test Registration scheme. Publishers can now apply to the Society to have their tests registered as 'tests'. The registration considers the evidence for minimum levels of psychometric adequacy and grants a certificate of registration to tests that meet these standards. The standards are based on the European Federation of Psychologists Associations criteria for test reviewing. In time, it is hoped that this registration will become widely known and recognised by test users and test takers alike as an indicator of quality.

Fairness

The user should check that appropriate steps have been taken in the development of the test to ensure fairness for applicants from different groups and backgrounds. Users should monitor the impact of the test in use and where there is any indication of adverse impact reassess the way the test is used (e.g. how cut-scores are determined), the test's validity and, indeed, whether it should be used at all. The use of high cut scores can exacerbate the adverse impact that will result from a measure if there is any difference in performance between groups and for this reason too they should be avoided.

Impact of cheating and distortion

If a proportion of candidates have managed to improve their scores through cheating or deception, the use of a top down selection rule will likely result in some of these people being selected for the next round of assessment at the expense of more suitable candidates who responded honestly. This would be unfair to those excluded and might leave the recruiter with insufficient good candidates at the next stage, and without verification checks of scores, an inappropriate candidate might be appointed.

Using a fixed cut-score where cheating has taken place does not exclude any well qualified candidates but may result in more people passing on to the next stage of selection than the recruiter would expect or ideally want. Assuming that the recruiter uses a follow-up validation test to identify and discard the cheaters a fixed cut-score will not result in any bias in final decisions even where distortion is relatively common. Thus, for reasons of both practicality and fairness, it is important to use pre-defined cut-scores that are based on a good rationale, even though this could lead to higher numbers of people passing than expected.

The candidate experience

Recruitment is a transaction between an employer and a candidate. The process, probably more than the outcome, will affect the candidate's perceptions of the employer. Image is important to employers who wish to attract the best candidates or for whom candidates are also potential customers. For this reason as well as the basic need to treat candidates with respect, it is important to consider how the candidate experiences the recruitment process. If it is a frustrating and bureaucratic process, if candidates are left in limbo with no information about what is happening, their perceptions will be negative. However if the system is easy to use and candidates understand the selection process, and receive timely and courteous communication about their progress, they are more likely to see it, and by inference, the employer, as fair. Access to feedback is helpful in the process, particularly where psychometric instruments are used. Applicants may appreciate learning what the instruments say about them, since this is not transparent in the way that other information provided on an application form is. Automated reports can be built into the system and provided there is someone a candidate can contact if they are worried about anything in the feedback, can be a cheap and effective way to deliver feedback.

Finally, genuine, honest candidates are unlikely to be put off by various checks and balances put into the system to detect cheating or collusion and to verify data. On the contrary, they are likely to see these as positive features of the system that will enable them to do well in a fair and just competition.

Conclusions

The internet is a fantastic new communication medium with great potential to handle the bureaucratic sides of recruitment. It can bring job openings to the notice of a broader range of people than a single job advert, it can make it easier for applicants to apply and it provides instant communication. It also allows assessments to be carried out at an earlier stage in the selection process, which can improve the validity of first-stage sifting and save time and money assessing less suitable candidates at a later stage. However, it is important not to get carried away with the excitement of its potential and forget the principles of best practice in recruitment and selection. The risks associated with remote administration need to be carefully managed if the potential value of this medium is to be realised.

Bibliography

The ITC Guidelines are available from the British Psychological Society Psychological Testing Centre (www.psychtesting.org.uk) and from the ITC website (www.intestcom.org)

American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing.* Washington DC: American Psychological Association.

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

International Test Commission (2001). International guidelines for test use. *International Journal of Testing, 1,* 93–114.

International Test Commission (2006). International guidelines on computer-based and internet-delivered testing. *International Journal of Testing, 6,* 143–171.

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