

Digital Assessment Industry Standards

Presented by the Association for Business Psychology

Points of Contact: Skye Trubov, Alan Bourne

These standards for digital assessment are meant to retain the scientific rigour in the design and delivery of assessments, irrespective of platform or delivery mechanism, to ensure a positive candidate and administrator experience.



Contents

04	1. Introduction
04	Purpose and rationale
04	Vision, goals for body of work
05	2. Definitions: what is a “digital assessment”
05	Context
05	Mobile assessment
06	Gamification and game-based assessments
06	Social media profiling
07	Virtual reality
07	Big data, machine learning and robotics
08	Digital methods: current, potential and unimagined
09	Current
09	Potential
10	Unimagined
10	Experience
10	Assessors
11	Participants/candidates
12	3. Ethical and legal context
12	Informed consent for both assessors and candidates
12	Managing confidentiality
12	Data protection and security
12	Employment law
13	4. Principles of Digital Assessment
13	Core principles that apply to all assessments, adapted for digital
13	Validity in digital contexts
16	Reliability in digital contexts
18	Digital standardisation
19	Fairness in digital contexts
21	Legal defensibility (known stipulations and general litigation)
21	United States
23	United Kingdom
23	European Union
23	Data protection and management
24	Experience
24	Consider the interaction of humans and technology:
24	Accessibility and equivalence
24	User engagement/experience
25	Solution design specifics (e.g item writing)
25	Assessor/administrator engagement/experience
25	Design software and ergonomics

Contents (continued)

25	Business benefit
25	Standards for different types of assessment that may be delivered digitally
26	Personality/behavioural
26	Cognitive
26	Motivation
26	Situational judgement
27	Biographical
27	Video and online interview
27	Game-based assessments
28	Online assessment centre exercises
28	Immersive/virtual simulations
28	360-degree feedback
29	Biometrics
29	Social media and other internet data
29	Personal data and records (e.g health)
30	References

1. Introduction

Our purpose for proposing a set of Digital Assessment Industry Standards (DAIS) is to retain the scientific rigour in the design and delivery of assessments, irrespective of platform or delivery mechanism, to ensure a positive candidate and administrator experience. There is a balance between rigour and user experience, so we seek to establish a set of standards and guidelines that are practically viable. Quality is not the only incentive though – ethical and legal concerns also drive the initiative.

Purpose and rationale

Advancements in technology continue to chart new methods for measurement, as the development of new platforms and delivery mechanisms advances at a rapid rate. The lack of guidance around digital assessments carries a huge risk for the implementation of such unregulated solutions. The psychology community is well positioned to offer standards of practice and to provide clarity for administrators and suppliers alike. Without standards, the industry may see unfounded claims go unchecked around assessment tools that lack rigour and validity, potentially damaging the reputation and professional work of occupational psychologists and companies that employ such practices. Additionally, ethical and legal implications can transpire without adequate rigour in assessments.

Psychological assessments must be built around certain scientific principles, but should also be practically viable so that good assessment use can be easily adopted. This initiative seeks to improve a fragmented process which has created vulnerability in the industry and ultimately, aims to improve practices and products in the marketplace.

Vision, goals for body of work

This will be a living document which is intended to be a scalable and relevant resource. The content within the document and its subsidiary pieces will be digestible and practical to ensure continued use. The information is delivered on an interactive platform for an engaging experience, encouraging an informed decision-making process.

2. Definitions: what is a “digital assessment”

Digital assessment is a term to describe the use of technology and the internet to measure and identify skills, abilities and work-related potential, often leveraging traditional selection methods and combining them with technological advancements.

Digital assessment is not a type of test or construct, it is a method of measurement. This means it can take a large variety of forms, so long as two key factors are satisfied: technology and measurement of skills, abilities or work-related potential. This makes digital assessment applicable to the technology platforms already in use (e.g. mobile, video-interviewing, gamification; see ‘Section b. Digital methods’), as well as new options (e.g. virtual reality, social media profiling).

Context

We live in a hyper-connected world where our real-world existence is very much intertwined with our digital selves, a knock-on effect facilitated by the rapid advent and uptake of new technology. More importantly, a shift in behaviour toward digital has had a real impact on how we work, due to the heavy dependence on technology. Technology solutions, robots and machines will continue to replace many of the jobs currently performed by humans in an ever-increasing demand for efficiency (Thompson, 2015). In light of this paradigm shift and other demands for workplace productivity, it is more important than ever to harness digital assessment in the search for talent.

The early stages of digital assessments surfaced along-side development of computing power and the speed, processing and stability of internet connectivity. Online platforms, facilitated by advances in technology, measurement theory and practice gave way to the first pen-and-paper psychometric and personality tests to be adapted for use on computers. Improvements such as these drive bottom-line impact for organisations and lead to further efficiencies regarding administration, scoring, reporting of tests in addition to data management and access to a wider pool of applicants (Scott & Lezotte, 2012). This is ideal for industries with high employee turnover such as the hospitality industry (Baum, 2007), since the tests are available anytime and anywhere, allowing for Unproctored Internet Testing (UIT). These online tests still make up a significant portion of current digital assessments, especially when it comes to high-volume recruitment for non-management positions (Reynolds & Rupp, 2010).

Mobile assessment

Mobile internet and devices have matured to the point where it is now practical to deliver digital assessments through mobile devices. GSMA data (2016) shows that more than 50% of browsing globally is via mobile devices as opposed to a desktop or laptop computer, and 20% of Millennials in the UK report as not owning a computer and only using their mobile to access the web. With a large proportion of people accessing the internet primarily through their phones, mobile-enabled digital assessments have become much more in demand. Moreover, because most other online services are already delivered in a mobile format, there is a growing expectation that everything will be mobile-enabled—online assessments included. Parallels can be drawn between the shift towards mobile assessment and the shift of internet-based psychometric assessment.

Gamification and game-based assessments

Gamification and game-based assessments (GBA's) use games and gaming features in non-game contexts for measurement. With a drive from companies toward candidate engagement and experience, digital assessments are likely to take on additional game-like features; these features can include the addition of leader-boards, avatars or badges to existing assessments. Several companies have taken the lead on GBA's and innovate at the forefront of providing fully gamified assessments. The boundary between an online game played for some challenging fun and a pre-employment assessment that influences hiring decisions is considerably blurry at present. Organisations may continue to introduce gamified assessments and GBA's as part of a formal application process, or potentially could use psychometric game badges to identify talent (Chamorro-Premuzic, Winsborough, Sherman & Hogan, 2016). It's possible that the market sees a divergence of utilisation.

Social media profiling

Social media activity generates significant amounts of data mostly in the public domain. Although many organisations have been using social media and other internet information to help make decisions about job applicants for several years (Stamper, 2010), there is limited empirical evidence to support this practice. For example, a study by Van Iddekinge, Lanivich, Roth and Junco (2016) revealed that recruiter ratings of Facebook profiles have zero correlation with job performance, turnover intentions and actual rates of turnover. Assessors seem attracted to social media as a platform for digital assessment—with or without due cause—because it is believed to demonstrate a more genuine image of the candidate and there is an allure to reap benefits from this large amount of available data (Miguel, 2013).

Many candidates expect their digital lives to be examined for hiring purposes (El Ouiridi, Segers, El Ouiridi & Pais, 2015), resulting in a great deal of time spent curating their online personae. This behaviour results in the same degree of impression management and social desirability as resume curation (Back et al., 2010). Current developments in social media profiling involve sophisticated machine-learning algorithms in hope of controlling for as many variables as possible. Algorithms are used to mine social media data and make predictions regarding a candidate's profile (e.g personality and interests), which tend to outperform human inferences (Lambiotte & Kosinski, 2014; Youyou, Kosinski & Stillwell, 2015). Companies like Entelo and Connectifier take advantage of these machine-learning algorithms and data mining techniques to provide businesses with insights about their active and even passive candidates. Viability of this method remains dubious as there are significant issues of informed consent that would need to be handled appropriately for these methods to become more widespread, alongside demonstration of their validation or contribution.

Virtual reality

Virtual Reality (VR) uses real-time computer graphics to create a computer-generated environment that looks and feels real to the user. Although the technology has existed for nearly 20 years, only now has it reached adequate quality to be used in the consumer market. As a result, VR has the potential to become a key platform in digital assessments as a technological evolution of computer-based simulations that assesses candidates. For example, Aguinis, Henle and Beaty (2001) illustrated potential applications of VR to assess skills and abilities. For example, communication skills can be assessed by analysing interactions with virtual subordinates, and psychomotor skills through manipulation of small virtual objects. Meanwhile, Cascio and Aguinis (2008) argue in “Staffing Twenty first century Organizations” that the modern staffing process needs to take context into account by ensuring the selection environment emulates the work environment. They suggest that there is a need for even more realistic and complex environments than those already created in assessment centres to allow for the simulation of hazardous tasks or rare occurrences— VR could be a solution for such a proposition, however not without concerns like cost and feasibility.

While the use of VR in the assessment space poses significant opportunity, this methodology has yet to be adopted by the mainstream marketplace. A few companies are in early stages of trialling this technology, so the industry can expect to see results by the end of 2018.

Big data, machine learning and robotics

Recent innovation has popularised tools designed to evaluate massive data sets from digital assessments. Big data on organisational performance can be used to reveal patterns, trends and associations which can then be used to develop models to create predictions. These predictions can then also inform digital assessments used for recruitment. Algorithms have even been applied to the evaluation of video interviews, where candidates record their answers to structured interviews, potentially eliminating the need for an assessor altogether. For example, automatic reading of voice-generated emotions, candidates’ word choice, speech patterns, facial expressions and pupil is used in video interviewing to inform hiring decisions. As such methods are subject to further research, the extent and limitation of their validity will become clearer along with their potential to both aid and automate assessment activity.

It is important to note that some of these more technologically advanced digital assessments have not come to the forefront without concerns from experts. Some key considerations:

1. There have been issues around people’s privacy and anonymity. How for instance is consent informed and how does an individual maintain ownership of how a potential employer makes use of their personal data, even if it is publicly available?
2. It is unclear whether the costs involved in developing digital assessments is always justified against any additional accuracy attained versus established tools, since few studies have attempted to establish this.

3. Even in an attempt to reduce bias and adverse impact, algorithms often reproduce the biases inherent in its creator, in the data it is using or in society at large (Sweeney, 2013; but see Feldman, Friedler, Moeller, Scheidegger & Venkatasubramanian, 2014 for a potential fix). This makes it especially hard to justify digital assessments carte blanche above traditional tools.
4. As recruitment processes may become more algorithm-based, there is a risk of less two-way interaction between a candidate and the organisation that is necessary to establish mutual expectations of future working relationships (psychological contract).

In addition to these concerns, there are some that view these technological advancements in digital assessment as a fad or another trend that will come and go (Tuzinski & Kantrowitz, 2016). Based on how much the industry is expanding, there is likely durability in some of the trends and technologies, yet much positive bluster still exists. The true value of these technological advancements remains to be demonstrated, and the proof will be in what outcomes can be shown versus the costs involved.

In the meantime, we still need to be critical of our assessment processes. As seen earlier, many of the technological advancements have already been adopted and used for selection, so these are the assessment processes that are here right now. Moreover, some papers have addressed this issue by providing specific guidelines regarding the research and shifts in perception necessary to prevent digital assessments from becoming a fad (Angrave, Charlwood, Kirkpatrick, Lawrence, & Stuart, 2016; Rasmussen & Ulrich, 2015).

Digital methods: current, potential and unimagined

This a comprehensive but non-exhaustive list of current and potential methods that have been identified as able to deliver and/or facilitate digital assessments.

- **Video interviews** – structured interviews that ask candidates to respond via a webcam or mobile phone camera or laptop/desktop camera to pre-recorded questions, images or work-samples using video chat software. Often, these are then rated by a person as in face-to-face structured interview. However, current advancements in this field include using algorithmic reading of voice-generated emotions, candidates' word choice, speech patterns, facial expressions and pupil dilation to inform hiring decisions.
- **Automatic resume screening** – a sifting software that automatically removes formatting from resumes, highlights keywords, sorts information accordingly and matches this against employer's list of desired skills and keywords. This often results in an aggregate score of sorts that determines progression to the next stage of recruitment.
- **Online psychometric assessments** – online and automated versions of previously pen-and-paper psychometric assessments, which offer a quantitative assessment of psychological attributes (e.g. verbal reasoning ability, personality characteristics). Development of such assessments so they can be delivered via mobile devices represents one aspect of change in this arena.

- **Simulations** – a measurement method that incorporates multimedia, such as audio, video and 3D animation, to provide a candidate with a representative work sample of what would be required on the job and to test higher-order skills (e.g. strategic thinking, decision-making). These can include desktop simulations, often in the form of e-tray exercises, virtual environment-based simulations and character/scenario-based simulations (Hawkes, 2013).
- **Gamification** – the use of game elements in non-game contexts (such as assessment) to enhance a candidate's experience, motivation and improve the validity of the assessment. This can range from simply adding graphics, storylines, leader-boards and badges to an existing process, to creating a stand-alone game that measures the desired constructs (Armstrong, Ferrell, Collmus & Landers, 2016).

Potential (and early stages)

These are emerging platforms that currently raise a significant number of questions and have a very limited adoption rate.

- **Virtual reality (VR)** – a computer technology that enables users to be immersed in an alternate world. Through the use of real-time computer graphics, users experience a computer-generated environment as if it is real and they are part of it (Aguinis, Henle & Beaty, 2001).
- **Augmented reality (AR)** – very similar to VR (see above), but instead of fully immersing the user, the technology superimposes computer-generated images on the user's view of the real world. This results in a composite view combining a virtual and a real world.
- **Social media profiling** – the review of publicly available online information from websites designed to connect individuals, such as Facebook and LinkedIn (Roth, Bobko, Van Iddekinge & Thatcher, 2016). This often involves a hiring manager simply searching for information on applicants and manually reviewing it. However, services exist that automatically collate all the necessary online information, profile it and even use machine learning to inform hiring decisions
- **Big data analytics and AI** – analysis of extremely large data sets on organisational performance to reveal patterns, trends and associations relating to human behaviour and interactions. This can be used to develop models to be used to create predictions. Linked to this, the development of artificial intelligence tools is likely over time to greatly increase the problem solving and analytical capacity available which may open up many new options for assessment.
- **Social robotics** – machines, often programmable by a computer, capable of carrying out a complex series of actions and judgements automatically, such as interviewing and shortlisting potential candidates.

Unimagined

When it comes to technology and the rate at which it has been developing, it can be safe to say that there are likely future platforms yet to be imagined. By comparing the list of current and potential methods, one may anticipate future approaches to be even more complex, even more realistic and to further address human bias and cheating concerns. As suggested by Chamorro-Premuzic et al. (2016), it may even be that theories are informed and re-evaluated based on current and upcoming platforms. The key point here is to recognise that technologies will continue to evolve and this will present both opportunities and issues for assessment.

Experience

Experience influences both those administering and taking the assessments; a scientifically rigorous assessment can and should still provide a positive interactive experience.

Assessors

Digital assessments have typically allowed for easier administration of assessments as they usually require no supervision and generate reports automatically, allowing an assessor to focus on interpretation of the analysis. However, the administration and communication with candidates must be done well, regardless of the method being used.

Scoring of digital assessment can also make a big difference in providing a simplified experience for assessors. Items and item types can easily be combined into scoring algorithms to assess candidates automatically. In addition to quantitative scoring as used in psychometrics, there have even been developments in automatic scoring of qualitative (i.e. open-ended) assessment responses and analysis of speech patterns from voice profiling. These and other measurements often provide additional data that would not otherwise be obtainable, potentially resulting in more consistent and higher quality scoring. Moreover, machine-learning algorithms can autocorrect in real-time, if necessary, which cannot be done with a static scoring criteria. This lessens the burden on active assessors and substantially reduces the margin of human error. Reporting is also increasingly automated, thereby benefitting the assessor greatly in terms of reduced demands on their time.

However, many of these experiences depend very heavily on thoughtful, scientific implementation of an assessment, as otherwise this can result in a vast array of issues. These can range from assessments that produce data in an undesirable format and other technical issues, to the use of algorithms that hinder selection or can embed bias or adverse impact unintentionally. Indeed, research suggests that recruiters are much more likely to adopt new technology in their selection methods, if they have high computer self-efficacy (Oostrom, van der Linden, Born, van der Molen, 2013).

Participants/candidates

Candidates increasingly expect assessments that are interesting, engaging, time efficient and delivered and evaluated at the touch of a button. Digital assessments make this increasingly possible. For example, it has been suggested that the popular trend for gamification in digital assessments enhances candidate's experience (Collmus & Landers, 2015). Digital assessments also often allow for branching and adaptive testing, meaning a candidate is challenged just the right amount without the assessment seeming unfair or too easy. Caveats do however, remain: Price and Patterson (2003) identified seven negative factors frequently associated with digital assessments – dehumanisation, lack of feedback, technical issues, attitude, lower candidate motivation, unfairness, and frustration.

Some early research on digital assessment concerned itself with some candidates having a poor experience due to anxiety about using computers (e.g., Brosnan, 1998; Mahar, Henderson, & Deane, 1997). However, the prevalence of technology in our world may mean that this is no longer an issue, with research largely suggesting that computer anxiety is simply an expression of general trait anxiety (Frericks, Ehrhart, and O'Connell, 2003). Indeed, for candidates who are digital natives, digital assessment feels more natural, than non-digital alternatives, as revealed by research on the effects of age in technology adaption (Renaud and Van Biljon, 2008).

Several studies have shown that perceived usefulness and perceived ease of use are important predictors to a candidate's reaction towards a new technology in personnel selection (Brenner, Ortner & Fay, 2016; King & He, 2006). Although the former has not received much research, ease of use has been an important consideration in digital assessments (see Bartram, 2005). Firstly, it is important that candidates receive a standardised experience, regardless of the browser, operating system, platform or screen size/resolution used. This is not to say that every candidate has to have an identical experience, but that it has to be equivalent. This may include accounting for accessibility and disability. Secondly, the readability and legibility of content must not be affected by the aforementioned factors. For example, consider whether charts and diagrams required for a task will fit on every screen and whether the nature of the task will be adversely affected, if they do not. Lastly, if speed is a factor in an assessment, it is also important to consider whether the platform will affect the ease of use and response input, including dexterity.

3. Ethical and legal context

The use of any assessment in both candidate selection and development must be appropriate in the context of ethical considerations and legal constraints. These are summarised below.

Informed consent for both assessors and candidates

Both candidates and assessors should be fully informed of what they are signing up for when they participate in an assessment process. Any data they provide, whether through completing assessments or as an assessor, providing observations of other people, should happen in the context of knowing how that information may be used. This principle of 'informed consent' underpins how assessment is used ethically, and it is an obligation of anyone providing assessments to ensure this is observed.

Given assessment is often provided by a third party to an organisation, it is also important that third party providers ensure that adequate briefing and training to ensure that organisations use such data ethically. This standard applies to traditional, current assessment but is just as important with digital assessment—especially where data can potentially be leveraged for other purposes beyond that for which it was originally provided.

Managing confidentiality

When individuals participate in an assessment, they are typically doing so with varying degrees of confidentiality. For instance, in selection contexts they can expect their data to only be used for the purpose they have agreed to (e.g. applying for a specific role) and for information to be viewed only by those who have a need to see it, such as recruiters or the hiring manager in question.

Data protection and security

In most markets, there are clear data protection laws in place that must be adhered to, giving close attention to how data is used and stored. It is incumbent on organisations to ensure that data is securely stored and not vulnerable to being accessed (for instance by computer hacking). These vary by geography and more detailed information is available here ([ADD LINK TO MORE DETAILED SECTION ON THIS TOPIC](#))

Employment law

In many countries, legislation exists not only around data protection but also the fair use of assessment information in employee selection decisions. For instance, this may relate to gender, ethnicity, age and a range of other protected characteristics dependent on the jurisdiction. It is incumbent on assessment users to ensure they are operating within these constraints, and ideally going above and beyond this to deliver fairness as effectively as possible by removing bias from assessment processes. In the digital world, this becomes increasingly easier to track for all assessments, be they psychometric, simulation or interview based.

4. Principles of Digital Assessment

This section offers guidance to aid decision-making around digital assessments.

Core principles that apply to all assessments, adapted for digital

As summarised in the previous section on 'psychological assessment', certain core principles pervade all psychological assessment, irrespective of delivery mechanism. New methods for measuring psychological constructs and individual differences should adhere to the following core principles so that rigour is not lost in the process, and to avoid damage or useless results:

1. Standardisation
2. Reliability
3. Validity/efficacy predicting outcomes
4. Fairness

Validity in digital contexts

Validity may be the most important principle to maintain in any psychological measurement. The critical question to be asked is whether an assessment is valid; in other words, can be shown to quantifiably predict relevant outcomes (e.g accurately identifying high performers).

In simple terms: does it work?

1. What evidence shows that the assessment **predicts performance/desired outcomes**?
2. Can the assessment predict success across function? Is it **generalisable**?
3. Is the assessment relevant to the context where it will be used?

A bottom line question when choosing a test

To check if the above requirements are met, the following question can be asked by the organisation:

If 'we' were challenged legally (e.g. in an employment tribunal) around why this test was chosen, are 'we' confident to explain this choice clearly in a way that is defensible?

These topics can be further understood by elucidating on a few key constructs.

Criterion validity

To determine criterion, the question to consider is:

Q1a: *Is there significant evidence that the assessment predicts performance for the role in question?*

The key aim in practical terms is to determine how frequently the assessment correctly identifies a high performer. Is this no better than chance (50%)? Or does the assessment accurately predict who is a high performer 75% or 80% of the time? This is often expressed in terms of a significant correlation between the results of the assessment and performance. Converting the statistics to an easy to understand number such as the percentage of people accurately identified as high performers can be very helpful to illustrate the impact. For instance, using numerical representation can be useful when communicating with stakeholders as numbers and graphical data more straightforward to understand.

The criterion validity (i.e. accurate prediction of on the job performance), is often expressed in statistical terms as a correlation, ranging from 0 to 1. The higher the better; an expectation of 0.3 or higher holds true for an individual assessment and ideally, the number is nearer to 0.4 or 0.5 for stronger validity.

This kind of analysis is typically undertaken through a validation study to link the performance of 100 or more individuals with the assessment results, to understand how one predicts the other. In practice getting a sample of this size isn't always feasible, which impacts on the statistical confidence of the findings.

If we have direct evidence that the assessment predicts performance for the role or context it will be applied, this quantifiable evidence is in many ways the best confirmation that the test works.

In situations where there isn't direct evidence available, or where it doesn't seem practicable to conduct a validation study, then the next option is to look for evidence of generalisability. This means evidence that the assessment has been validated elsewhere in a similar context which gives reasonable confidence that the assessment may be valid for the purpose in question.

If the first question has not been met, the second question to consider is:

Q1b: *Is there sufficient evidence of the generalisability of the assessment in other roles or contexts?*

- Many assessments, for instance cognitive ability and personality questionnaires, have a good degree of generalisability across roles. The following points serve as reference on how to determine this characteristic of an assessment:
- Review relevant studies or manuals from the supplier that show the validity of the assessment in other contexts. This should involve criterion validity studies where quantified statistical links have been made.
- Evidence should clearly show the correlation between the assessment and performance in the role, which may also be quantified in terms of the business impact. A local validation study should give significant correlations with relevant performance criteria as well.

In terms of the correlations that might be quoted, the higher the better and ideally you might expect to see 0.3 or higher for an individual assessment and ideally nearer 0.4 to 0.5.

- There should be adequate evidence of the relevance of the attributes measured to different types of jobs.

Overall, the body of evidence should support the view that because there is validation of the assessment for roles like those which need to be filled. Then, there can be some confidence that the intended assessment will be valid for the population where you intend to use it.

In summary, in the absence of an affirmative answer to question 1a, a solid answer to question 1b will still provide confidence that an assessment should, in general, be a valid predictor of relevant outcomes such as performance.

Given the above, it is also important to give some consideration to the specific role and circumstances for an assessment. What are the key competencies or behaviours that matter, based on the role requirements or other relevant information about 'what good looks like'?

The key point here is to understand how relevant the assessment is to the context where it's used, considering how valid the content is for the situation.

Relevance

When examining whether a test is relevant or not, the question to consider is:

Is there evidence that the attributes measured by the assessment are relevant for the specific role or purpose in mind?

Whilst there may be evidence that an assessment is fairly generalisable (for instance numerical and verbal reasoning tests), it is important to be confident that they are relevant for the role or purpose where in use.

- For instance, numerical reasoning may be a highly relevant ability for graduate entrants into a financial services organisation, but is this type of test as relevant for IT developers or would an assessment of logical/inductive reasoning that is more closely linked to that type of role be a much better choice?
- Ideally, job analysis would provide evidence to identify the critical behaviours or a profile of skills that matter in the role, including competencies required.
- As a baseline, the job description, person specification and/or competency framework for the role should read across clearly with the attributes being assessed.

The above evidence will give confidence that what is being measured is relevant to the role in question. Combined with clear criterion validity evidence (ideally local validity, or at least generalised validity), this will provide a good basis for concluding the assessment is a genuinely valid predictor.

Reliability in digital contexts

Some methods of digital assessment will hold varying measurements of reliability, many will not hold true to traditional statistical methods of reliability (e.g Cronbach's Alpha) because response to certain assessments, particularly those which involve judgement, will change over time.

The bottom line question here is:

Is the assessment reliable? Are the results stable for each person assessed?

The reliability of an assessment relates to whether it can predictably measure the psychological characteristics it is intended to. This is not the same as validity which relates to whether the assessment predicts outcomes that matter, such as performance.

Reliability fundamentally relates to the stability of the measurement itself and how accurately it might measure a construct even over time. For instance, will the assessment measure the same characteristic today as it will in two weeks? This can be measured by looking for instance at 'test-retest' reliability – how correlated results are for a group of individuals across two points in time.

Since asking a sample of candidates to take an assessment twice is not always that practicable, an alternative measure often used for psychometric tools involves a statistic called Cronbach's Alpha, which looks at how closely associated the answers to different questions around the same topic tend to be. For instance, on average do people that tend to answer one question right also tend to get the other questions right? This statistic is widely used by psychologists as it's easy to generate, but is a surrogate – it doesn't actually measure reliability over time as such. The downside is that good 'Alpha' statistics can be achieved by adding lots of questions, which can make questionnaires longer and more repetitive than might be ideal from a candidate's perspective.

On more advanced assessments, which draw questions from a large database of items (i.e a bank), there is an alternative Omega statistic which is similar in concept to the Chronbach's alpha but works in this context where candidates may each get different questions to complete.

On all the statistics above (test-retest, Alpha, Omega), the statistics run from 0 to 1 and the higher the better. A statistic of 0.7 or more is generally considered a reasonable rule of thumb to give suitably accurate measurement.

Finally, on observed assessments, such as in a video interview or digital assessment centre, the equivalent measure is 'inter-rater reliability'. In this regard, an assessment should give adequate instructions and guidance such that suitably trained assessors will tend to arrive at similar judgements on a consistent basis.

One practical outcome of reliability is the accuracy with which results between candidates can be differentiated. In people measurement, every observation has a degree of inaccuracy around it. A score on a test on a given day should give a good measure of a candidate but it may not be exactly the same as the score they would get if they did it two weeks later. Within assessment, we are looking for enough accuracy that we can differentiate between candidates, but need to accept there is a **band of error** around every result. The higher the reliability, the smaller the band of error.

Whichever assessment process may be under consideration, there should be suitable evidence that the approach will give consistent and reliable results. These statistical tools mentioned are the way in which this evidence is presented.

A final note on this topic, is that validity is contingent on reliability. So, if there is a good evidence of validity, by definition you also have evidence that there is underlying reliability. This yardstick is often used with bespoke assessments like situational judgement tests, which don't tend to show high internal consistency but can be both reliable over time and retain good validity.

The bottom line question here is:

Is the level of reliability adequate enough to differentiate effectively between candidates in a meaningful and way over time?

Digital standardisation

Having established that an assessment solution should be able to provide a valid and reliable measure, the next consideration relates to comparing candidates to a suitable benchmark in a meaningful way.

In practice, most digital assessments of people involve some kind of benchmark. For psychometric tools, these are known as **norms** and often involve a sample of results from people who have completed the assessment.

In interviews, this is typically independent of comparing to other people and instead is against a clear **criterion** of what a score of '4' requires versus a score of '3' on a 1-5 scale.

Criterion-based assessments may have a minimum level or pass mark based on the desired level of capability, and this can be set by the organisation in relation to their needs.

For norm-based assessments (such as cognitive reasoning tests or personality questionnaires), people are typically rated by comparison to sample group. There are three basic ways of doing this:

1. Comparison to a large general norm group: the upside of this method is there may be many 100s or 1000s of people in the comparison group, and the downside is this group may be very general in nature.
2. Comparison to a more relevant group: for example, a norm for banking sector or a norm for graduates.
3. Comparison to a localised norm group: this is typically a group of current employees so gives an entirely localised benchmark and makes the norms as relevant as possible for the context in which it's being used.

Different measurement types are also more or less dependent on norms:

- For cognitive ability, performance can vary between role levels so picking the right norm is important to ensure it is relevant.
- In terms of personality, this varies mainly between people and less between organisations, sectors and levels. A localised norm can be comforting but may not vary as much from a general norm group.
- For situational judgement tests, because they are typically organisation-specific, the norms by definition often have to be very localised using existing employees as the situational questions would only make sense in that organisation.

The bottom line question here is:

Is there a suitable comparison group to accurately identify who is above, near or below the average for the role or group where the assessment will be used?

Fairness in digital contexts

Just as with any psychological assessment, the method, mechanism and platform must offer a fair experience to all participants. As digital assessments of different types tend to mean data is more thoroughly recorded (e.g. video interviews versus traditional paper-based interviews), there is much greater opportunity and expectation that the available data to test assessments are fair.

In this regard, there is also a legal context which varies between different countries (covered in the next section).

There is a widespread awareness now that decisions made in the recruitment process can cause **adverse impact**, which leads to proportionately fewer candidates get through the process on account of gender, ethnicity, age and other characteristics. A test that imposes adverse impact is not fit for purpose because it inherently draws an unfair bias into the process.

The legal burden of proof in this regard typically (e.g. in U.K. and U.S.) leaves some reasonable space for error, and acts on what has become known as the **four-fifth's rule**. This means that for every candidate from the advantaged group that gets through a selection process, at least 4/5ths of the disadvantaged group should also pass. If the proportion is lower, then adverse impact may be at a point where it is significant enough that if a candidate were to take a case to tribunal, they could argue successfully they were indirectly discriminated against.

Whilst many processes are 'fair enough' to be within the legal burden of proof, this is still quite a distance from achieving 100% fairness and inclusivity.

Essentially two main causes of unfairness appear in assessment processes:

1. Systematic bias in psychometric tools, for example cognitive testing.
2. Unconscious bias in interview and assessment exercise decisions

Fairness in psychometric tools

Looking at psychometric assessments, it is well known that cognitive tests like numerical reasoning can show average performance differences in terms of gender and ethnicity.

Cut-offs are often kept low on these tests to ensure the impact is within legal limits, however, we do still find systematic differences between groups around these abilities. Typically, a cut-off point at the 40th percentile, where no more than 40 percent of candidates are rejected, is a good rule of thumb in the UK for avoiding adverse impact. In the US, where differences between groups are larger (linked to the socio-economic context), the safe cut-off legally may be nearer the 30th or even 20th percentile on this type of assessment.

Part of the problem around fairness for cognitive reasoning tests in particular relates to how they are being used as a screen. The competencies needed to be successful are not just cognitive problem-solving that matters, but may also include resilience, empathy, leadership skills, etc. However, organisations often just sift on one aspect – cognitive ability – rather than taking a rounded view.

While there is a range of research evidence showing that cognitive ability tests have good generalised validity, often organisations screen on criteria that can cause adverse impact or limit how many candidates can be deselected at a given selection stage. This tends to be much less of an issue with personality or behavioural measures, and situational judgement tests are typically very fair in terms of low adverse impact.

One option widely used in markets like the US and elsewhere is blended assessment, which focuses on the range of competencies you need in your candidates, combining situational judgement, behavioural style and cognitive elements in one assessment. Essentially, candidates are assessed in the round on a number of balanced criteria. Whilst a particular group may on average score a little lower on a couple of competencies, they may score higher on others so these affects can average out. This approach can help mitigate adverse impact compared to standalone cognitive tests.

Fairness in interviews and assessment centres

The second major issue is the effect of unconscious bias in interviews and assessment centres. As these activities are often paper-based, the results tend to receive little scrutiny, however, they can also be a significant source of adverse impact due to unconscious bias.

For instance, an interviewer may make assumptions of someone 'looking right for the organisation' or they had been at a previous employer or university which is 'a good place to have been'. These assumptions are usually unconscious short-cuts people use to make decisions, but can drive decisions which are not inclusive.

What can be done to remove unconscious bias from interviews and assessment centres?

Organisations can offer recruiter and hiring manager skills and training. Training focused on this content can be an effective way to raise awareness and improve how judgements are made. In addition, organisations moving from paper-based to digital assessment centres or interviews have a great opportunity to use the data they capture to identify where the inclusiveness of assessor judgements can be improved.

The bottom line questions here are:

- *In relation to psychometric tools, is the content being used and the cut-off for decisions within legal limits relating to adverse impact?*
- *For observational measures like video interviews and assessment exercises, is there confidence that any bias will be picked up and are assessors adequately trained to avoid bias?*
- *Finally, raising the bar beyond this, how close to parity across groups is the decision-making overall? Does this meet the organisation's aspirations of fairness and diversity?*

Legal defensibility (known stipulations and general litigation)

Legal defensibility takes priority with all companies, especially global corporations.

United States

United States litigation outlines clear expectations for workplace conduct which acts within and in accordance to the U.S. equal opportunities/diversity legislation.

- Disparate impact/adverse impact is the primary legal issue:
- Involves an assessment indirectly leading to members of a protected class or group having less chance of being selected.
- Not about direct discrimination (occurs when someone has been 'singled out' due to their characteristics such as gender, race, age, etc.).
- Typically indirect discrimination from testing is due to systematic differences in group performance.
- Fundamental issue is different levels of access and support in relation to quality of education received and socio-economic factors.
- Cultural emphasis on education may also be a factor and can be reflected in group test results.

The questions to consider here are:

- *Does an assessment cause disparate impact?*
- *Is the procedure job-related and consistent with business necessity?*
- *Is there a less discriminatory alternative?*

Relevant legislation in the U.S regarding the workplace:

- Title VII of the Civil Rights Act of 1964
- Title I of the Americans with Disabilities Act (ADA)
- Age Discrimination in Employment Act (ADEA)

Need to demonstrate validity and fairness:

- 4/5th's rule in relation to impact on protected classes
- Risk of class action

When documenting analysis of assessments and assessment procedures, take into account:

- Type of assessment used and why it was chosen.
- Methods - sample and procedure.
- Descriptive statistics for each sub-group in the analysis.
- Include effect size analysis (SD difference between groups).
- Application of 4/5ths rule to assessment scores or overall role profile, split by each relevant group comparison.
- Clear recommendations for fair application of the assessments.
- Note that if a difference is not statistically significant at 5% level, it doesn't mean it can be ignored (comment on small sample sizes though would be relevant).

Summary:

- Any evidence of disparate impact creates a risk of lawsuit and potentially class action.
- Lack of evidence that a procedure is valid significantly weakens the case for its use and heightens the legal risk.
- Lack of proactivity in validating the applicability of an assessment and seeking to minimise disparate impact also heightens risk.
- The number of legal cases is significant but not huge.
- The appropriate analysis will also help achieve better prediction, enhanced fairness and greater business benefit.

Sifting and setting cut-offs:

Once we have completed our analysis for disparate impact, how can we then apply the findings in practice?

- The key is to set the selection 'sift out' or 'cut-off' points based on the results.
- Cut-off points need to be both practically useful (sift out candidates) but also fair (hence avoiding legal risk).
- They should also be backed up by validation evidence, as the underpinning of why the assessment is being used.
- High cut-offs disproportionately introduce adverse impact.

United Kingdom

The validity of the assessment is not a primary issue in UK equal opportunities legislation. It is legal, for instance, to randomly deselect applicants ('toss of the coin') when these are in very high volume or for instance in redundancy situations when other criteria are equal.

Three key steps to consider are as follows:

1. Does an assessment generate adverse impact to a particular group in a systematic way (i.e. on average to candidates from that group)?
2. Does the claimant belong to that group?
3. Was the claimant affected by the adverse impact in this specific claim?

European Union

EU workplace conduct varies.

Data protection and management

Strict laws already govern data management and protection, and new laws may be introduced in the future (for strengthening or weakening) as the web continues to facilitate the proliferation of data exchange. The key point is to be aware of and maintain up-to-date knowledge on what can and cannot be shared, extracted or used for assessment purposes.

Experience

Experience of the candidate and assessor depends on the type of assessment and how that assessment is administered or delivered.

Consider the interaction of humans and technology:

Humans tend to resist change as a psychological tactic to help make sense of the world. Digital assessments may be viewed as a change, so may be met with resistance. Managing this resistance on both sides – candidate and assessor – can be mitigated by making the right decision in choice of assessment.

Regardless of purpose (e.g recruitment, leadership development), an assessment that fits with the ethos, culture and values of an organisation will likely align with a particular vision or end-goal. Being able to explain in plain, real terms as to why the organisation is making use of a new type of assessment (i.e implementing change), is crucial for progress and successful adoption of the change.

Accessibility and equivalence

For an assessment to be viable for a generalised population, it must have the flexibility to accommodate those with physical and learning disabilities to reduce unfairness and bias.

User engagement/experience

User experience has a muddled definition, but can refer to the non-utilitarian components of a technology that mostly manifest in emotion, sensation, meaning, value and perception from interacting with a technology or platform (Law et al., 20014). The user experience is particularly relevant for overall satisfaction and engagement with a process and more substantially, with a brand.

A critical component of any assessment is experience; however, the user interface and user experience cannot supersede scientific rigour lest the product fail to be considered an assessment. A balance is critical: it is important to provide a positive experience to the candidate by introducing relevant methods of assessment in recruitment, but these novel platforms must retain key scientific principles to be valid measures of psychological constructs.

In short, a digital experience must present certain levels of validity and reliability to be considered a viable assessment – if not, the tool may be engaging but cannot qualify as an assessment mechanism. This is not to say that a digital platform can't offer an engaging experience for a candidate or attract breadth of talent, but it cannot be considered a psychometric assessment from which decisions on hiring can be made.

Another critical consideration for user experience is what the outcome and impact may be if the platform might crash, glitch or fail. In practical terms, if a candidate participates in an assessment on a VR platform, consider possible physiological and psychological consequences of the technology (e.g vertigo) and potential complications therein (e.g disorientation from a defect). A question to ask may be similar to: what happens then the candidate takes off the headset? These are important considerations of digital assessment that specifically require not just scientific rigour with regards to the assessment itself, but also with regard to a psychological appreciation of the candidate experience.

Solution design specifics (e.g item writing)

Items used in a digital assessment can be a challenge to write, depending on the platform. For example, existing numerical reasoning items that incorporate graphs and charts will not likely translate well into a small screen of a mobile device. As long as traditional psychometric items prevail, transferring pre-existing analogue (or digital) questions onto a new digital platform may or may not be feasible. User experience influences the viability of item writing and format because platform and delivery must be accounted for when validating an assessment.

The broader question is, can/should pre-existing analogue items transfer into a digital format in the first place? A concern in academia around the digital landscape points to the loss of validity when transferring items such that an existing test requires re-validation if placed onto a new platform.

A further consideration eliminates 'items' as they exist in current and known formats and extends measurement beyond what is considered an item. In social media profiling, 'items' would not be presented to candidates in the same way as verbal reasoning items would be – using data and other methods of gathering information is altogether new. Such measurement requires novel methodology for validation exercises.

Assessor/administrator engagement/experience

Similar to value around the candidate experience, it is important for an assessor to have a positive experience as well. This in turn supports a good candidate experience (depending on levels of involvement or proctoring).

Design software and ergonomics

Software design is paramount for a positive user interface, user experience, functionality and stability for digital platforms that run on a server. Data privacy and storage are key concerns: who has access to the data? Where is it held? What are the liabilities? If outsourcing, what are the legal components?

Business benefit

There is no question that businesses in each category – corporate and supplier – can profit from a well-designed digital assessment. Technological advancements allow for increased efficiency for current and potential assessment delivery mechanisms, but also come with consideration over potential negative ramifications.

Standards for different types of assessment that may be delivered digitally

The following summary highlights some of the key considerations to bear in mind with different types of technique when they are applied in a digital setting.

Personality/behavioural

Can the results of the assessment be linked to meaningful outcomes for the situation where it will be used, such as performance in the role? Has this been demonstrated statistically?

Does the format of the assessment work across different device types (e.g. mobile, desktop) in an equivalent fashion?

Is the assessment an effective candidate experience, avoiding unnecessary repetitiveness which provide reliable measurement across a range of personality / behavioural scales?

Cognitive

Is there evidence that the assessment is a valid and reliable predictor of performance in the role or roles in question?

Has the potential for adverse impact been reviewed and where possible mitigated, or relevant safeguards (such as maximum cut-offs) suggested?

Does the format of questions provide equivalence across device types (e.g. mobile, tablet, desktop), so there is no significant advantage in using one device over another? Specifically, if questions are timed, does this operate similarly across devices?

Motivation

Does the assessment provide a valid method for understanding motivations and linking these to meaningful outcomes such as fit to an organisation or role type?

Does the format of questions provide equivalence across device types (e.g. mobile, tablet, desktop), so there is no significant advantage in using one device over another?

Situational judgement

Have the questions been linked to the role in question? Has validity been shown within the specific situation where the situational judgment questions will be applied?

Do the questions give a suitably realistic view of decisions required in the role?

Can the questions be presented legibly, accounting for accessibility and equivalence factors?

Biographical

Are any biographical questions used free from bias and fair?

If questions relate to experience, have this been reviewed to ensure they are free from the risk of age discrimination in particular?

What is the purpose for collecting the data, specifically? Can questions be clearly linked to relevant qualifications or performance in the role?

Video and online interview

Are the questions used clearly linked to competencies for the relevant role?
Have these been shown to be valid predictors of future success in the role?

Are there processes in place to ensure adequate reliability between raters is achieved?
How is this monitored?

To what extent has the potential for unconscious bias and adverse impact been monitored?
What processes are in place to ensure this is minimised?

Where automated scoring processes are in place, have these been shown to be valid for the role or roles and organisation where the process is being used? What is the validity and reliability?

What is the candidate experience like, and how might this effect the outcome?

Game-based assessments

What is the game-based assessment measuring and does it relate to key capabilities needed in the role?

Is there evidence that the assessment is a valid and reliable predictor of performance in the role or roles in question?

Does the assessment provide suitable opportunities for practice, to ensure individuals with less gaming experience are not disadvantaged?

Is there evidence to show that adverse impact has been reviewed and managed appropriately to ensure fairness?

Online assessment centre exercises

Are the exercises used clearly linked to context and competencies for the relevant role?
Is there evidence they are valid predictors of future success in the role?

To what extent has the potential for unconscious bias, systematised bias and adverse impact been monitored? What processes are in place to ensure this is minimised?

Where automated scoring processes are in place, have these been shown to be valid for the role or roles and organisation where the process is being used?

Immersive/virtual simulations

What is the simulation measuring and does it relate to key capabilities needed in the role?
For instance, is it a work sample reflecting real world activities in a role?

Is the simulation clearly linked to context and competencies for the relevant role? Is there evidence it is a valid predictor of future success in the role?

Does the assessment provide suitable opportunities for practice, to ensure individuals with less experience are not disadvantaged?

Is there evidence to show that adverse impact has been reviewed and managed appropriately to ensure fairness?

What is the psychological impact of the experience?

360-degree feedback

Are the questions within the 360-degree feedback process clearly linked to the competencies for the role?

Are there clear behavioural descriptions in place to ensure consistent interpretation and adequate reliability between raters?

To what extent has the potential for unconscious bias and adverse impact been monitored? What processes are in place to ensure this is minimised?

Biometrics

What is the biometric assessment measuring? Does it relate to key capabilities needed in the role?

Is the assessment a valid predictor of future success in the role? What additional objective data is it offering that links to performance?

Is there evidence to show that adverse impact has been reviewed and managed appropriately to ensure fairness?

Has genuinely informed consent been obtained from participants for the use of biometric measurement as an input to an assessment process?

Social media and other internet data

What is any profiling from social media responses (e.g. Facebook likes or similar) measuring? Does it relate to key capabilities needed in the role?

What is the reliability of the social media profiling?

Is the assessment a valid predictor of future success in the role?

Is there evidence to show that adverse impact has been reviewed and managed appropriately to ensure fairness?

Has genuinely informed consent been obtained from participants for the use of their data for assessment purposes?

Given social media use may have happened without expecting the data to be used for assessment, does the participant retain the genuine right to stop their data being used in this way?

What data has been used to design the algorithm? How can the methodology be free from bias?

Personal data and records (e.g health)

What is use of personal data such as health records required for? Is there is a clear link to conduct or performance in the job? (e.g. criminal records checks, or operating in a substance-free environment)

Is the assessment a valid predictor of performance or compliance with safety requirements in the role?

Has any requirement to share personal data been made clear prior to application? Has genuinely informed consent been obtained from participants for the use of their data for these purposes?

Is there evidence to show that adverse impact has been reviewed and managed appropriately to ensure fairness?

References

- Aguinas, H., Henle, C. A., & Beaty Jr, J. C. (2001).** Virtual reality technology: A new tool for personnel selection. *International Journal of Selection and Assessment*, 9(1-2), 70-83.
- Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016).** HR and analytics: Why HR is set to fail the big data challenge. *Human Resource Management Journal*, 26(1), 1-11.
- Armstrong, M., Ferrell, J., Collmus, A., & Landers, R. (2016).** Correcting Misconceptions About Gamification of Assessment: More Than SJTs and Badges. *Industrial and Organizational Psychology*, 9(3), 671-677.
- Back, M.D., Stopfer, J.M., Vazire, S., Gaddis, S., Schmukle, S.C., Egloff, B., & Gosling, S. D. (2010).** Facebook profiles reflect actual personality, not self-idealization. *Psychological Science: A Journal of the American Psychological Society/APS*, 21(3), 372–374.
- Bartram, D. (2005).** The Great Eight competencies: A criterion-centric approach to validation. *The Journal of Applied Psychology*, 90(6), 1185–1203.
- Baum, T. (2007).** Human resources in tourism: still waiting for change. *Tourism Management*. 28(6), 1383-1399.
- Brenner, F. S., Ortner, T. M., & Fay, D. (2016).** Asynchronous video interviewing as a new technology in personnel selection: The applicant's point of view. *Frontiers in Psychology*, 7, 863.
- Brosnan, M. J. (1998).** The impact of computer anxiety and self-efficacy upon performance. *Journal of Computer Assisted Learning*, 14, 223–234.
- Cascio, W. F., & Aguinis, H. (2008).** Staffing twenty-first-century organizations. *The Academy of Management Annals*, 2(1), 133-165.
- Chamorro-Premuzic, T., Winsborough, D., Sherman, R., & Hogan, R. (2016).** New Talent Signals: Shiny New Objects or a Brave New World? *Industrial and Organizational Psychology*, 9(3), 621-640.
- Collmus, A. B., & Landers, R. N. (2015, April).** Game narrative in personality assessment: The development of a scale. Presented at Virginia Psychological Association's Spring Convention and Education Conference, Virginia Beach, VA.
- El Ouiridi, M., Segers, J., El Ouiridi, A., & Pais, I. (2015).** Predictors of job seekers' self-disclosure on social media. *Computers in Human Behavior*, 53, 1-12.
- Feldman, M., Friedler, S., Moeller, J., Scheidegger, C., & Venkatasubramanian, S. (2015).** Certifying and Removing Disparate Impact. *Proceedings Of The 21Th ACM SIGKDD International Conference On Knowledge Discovery And Data Mining - KDD '15*.
- Frericks, L., Ehrhart, K. H., & O'Connell, M. S. (2003, April).** Computer anxiety and test performance: Comparing selection test formats. Paper presented at the 18th annual conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

References

- Hawkes, B. (2013).** Simulation Technologies. In M. Fetzner & K. Tuzinski (Eds.), *Simulations for Personnel Selection* (pp. 63-82). London, UK: Springer.
- King, W. R., & He, J. (2006).** A meta-analysis of the technology acceptance model. *Information & Management*, 43(6), 740-755.
- Lambiotte, R., & Kosinski, M. (2014).** Tracking the digital footprints of personality. *Proceedings of the IEEE*, 102(12), 1934-1939.
- Lezotte, D. V., & Scott, J. C. (2012).** Web-Based Assessments. In N. Schmitt (Eds.), *The Oxford Handbook of Personnel Assessment and Selection* (pp. 485-513). Oxford, UK: Oxford University Press.
- Mahar, D., Henderson, R., & Deane, F. (1997).** The effects of computer anxiety, state anxiety, and computer experience on users' performance of computer based tasks. *Journal of Personality and Individual Differences*, 22, 683-692.
- Miguel, R. F. (2013).** LinkedIn for hiring decisions: A content validity framework. In R. F. Miguel (Chair), *The promise and perils of social media data for selection*. Symposium presented at the Society for Industrial and Organizational Psychology, Houston.
- Oostrom, J. K., van der Linden, D., Born, M. P., & van der Molen, Henk T. (2013).** New technology in personnel selection: How recruiter characteristics affect the adoption of new selection technology. *Computers in Human Behavior*, 29(6), 2404-2415.
- Price, R.E. & Patterson, F. (2003).** Online application forms: psychological impact on applicants and implications for recruiters. *Selection and Development Review*, 19(2), 12-19.
- Rasmussen, T., & Ulrich, D. (2015).** Learning from practice: How HR analytics avoids being a management fad. *Organizational Dynamics*, 44(3), 236-242.
- Renaud, K. & van Biljon, J. (2008).** Predicting technology acceptance and adoption by the elderly. *Proceedings Of The 2008 Annual Research Conference Of The South African Institute Of Computer Scientists And Information Technologists On IT Research In Developing Countries Riding The Wave Of Technology - SAICSIT '08*.
- Reynolds, D. H., & Rupp, D. E. (2010).** Advances in technology-facilitated assessment. In J. C. Scott & D. H. Reynolds (Eds.), *Handbook of workplace assessment: Evidence-based practices for selecting and developing organizational talent* (pp. 609-641). San Francisco, CA: Jossey-Bass.
- Roth, P. L., Bobko, P., Van Iddekinge, C. H., & Thatcher, J. B. (2016).** Social media in employee-selection-related decisions: A research agenda for uncharted territory. *Journal of Management*, 42(1), 269-298.
- Stamper, C. (2010).** Common mistakes companies make using social media tools in recruiting efforts. *CMA Management*, 84(2), 12-14.

Sweeney, L. (2013). Discrimination in Online Ad Delivery. *Queue*, 11(3), 10.

Thompson, D. (2015, July/August). A world without work. *The Atlantic*. Retrieved from <http://www.theatlantic.com/magazine/archive/2015/07/world-without-work/395294/>

Tuzinski, K., & Kantrowitz, T. (2016). Meet You at the Peak: How I-Os Should Prepare for New Technologies. *Industrial and Organizational Psychology*, 9(3), 655-660.

Van Iddekinge, C. H., Lanivich, S. E., Roth, P. L., & Junco, E. (2016). Social media for selection? validity and adverse impact potential of a facebook-based assessment. *Journal of Management*, 42(7), 1811-1835.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112(4), 1036-1040.

www.theabp.org.uk

© Association for Business Psychology 2017

