



Current Development in PSYCHOMETRIC TESTS

BY ADRIAN FURNHAM

Finding the right person who is well suited for a job position is vital for any organisation's success. There are numbers of factors in which employers base their judgments and decisions; some of which are the personality, attitude and knowledge of the applicants. Test taking plays a vital role in knowing and understanding these aspects in a deeper sense. This article explains the importance of psychometric tests and its development over the years.

The enthusiasm for using personality and ability tests in selection and development waxes and wanes. It is still a big industry. There are many defenders and distractors who periodically “have a go at each other” on a number of repeated issues. What is the future of psychometric tests?

It is unusual to go on any training course without having to complete a test that tells you about your style, personality, or preferred team role. These are tests of attitudes and beliefs, of preferences and traits, strengths and agility, learning style, thinking style and coping style. These are tests of real general intelligence, of emotional intelligence, of practical intelligence and of multiple intelligence. These are tests of bright and dark side personality: normal as well as abnormal functioning. There are tests of motivation and values. Consultants and test publishers know that this is big business.



Test taking can also be a drug for a few narcissists. For most people, it soon loses its power because it is as much confusing as entertaining; as much befuddling as enlightening.

There are many reasons why trainers like tests. Tests can give people new concepts, language and hopefully self-insight even if they reject what the test says.

Tests give comparative, specific and numerical data. Unlike references and testimonials, which flourish on euphemisms and double speak, tests give numeric data. One cannot get away with “satisfactory”, “high spirited” and “quirky” in tests. Good tests have normative data so we can measure up oneself against population norms.

Test taking can also be a drug for a few narcissists. For most people, it soon loses its power because it is as much confusing as entertaining; as much befuddling as enlightening. If you are serious about using tests to improve your insight, choose one that has been psychometrically validated based on a social theory; and one that describes process, not only typologies.

The accurate and reliable assessment of people at work is essential for many decisions around selection, development, promotion and redundancy. It is a complex area but one that is clearly related to the health and success of any organisation.

Assessment involves a cost-benefit analysis for recruitment, retention and development. A good analysis requires well-defined and measurable objectives. Each assessment instrument or process should evaluate the degree to which objectives have been met. Furthermore, desirable and undesirable criteria

should be delineated so people can be “selected out” and “selected in” as well.

Many assessment tools have strengths and weaknesses, and are appropriate for measuring different objectives. However, some tools and techniques are superior to the others. Psychometric evidence provides a strong foundation for comparing different tests and for matching the right test or assessment with a particular objective.

Testing and assessment is a specialised field and may require an expert to ensure the test is used appropriately and effectively. There remains a lot of ignorance and myths about assessment methods which can have significant problems for HR managers in particular organisations.

Having researched this area for 30 years, evaluating the data, here is my considered response to 12 myths about testing:

1 All candidates fake and lie making the data worthless.

If everyone faked the good/ideal answer, they would all be the same and tests would have no validity. Clearly, some tests are more fakeable than others. You can not fake ability tests (only to do worse). There are a number of techniques that test constructors have for catching those who fake, including lie scales and obtaining templates of those

who they have deliberately asked to fake so that they know what a faking profile looks like. This is considered a serious problem by most people who somehow believe that people do not fake in interviews or on application forms, despite the evidence.

2 Testing is simply too costly in terms of time and money.

A surprising number of tests are free. If the cost of testing is taken into consideration compared to a candidate’s annual salary or the cost of failure and derailment, it is clearly very little. Some tests can be too expensive for what they provide, but the majority have a very good value, certainly if they seriously improve decision-making.

3 Tests are too unreliable: mood, health and the situation influence the results. In fact, the opposite is true.

Tests are surprisingly robust, yielding very similar results on different occasions, a long time apart. They are just as reliable as most medical tests and much more so than some (blood pressure measures). The circumstances that are likely to yield the most unreliable results is where a person is doing an ability test while feeling very unwell and taking the test in a noisy environment. Indeed, we have data of people taking the same test 50 years apart with surprisingly similar results.

4 Tests do not predict work performance.

This is perhaps the most important issue: what is the relationship between test scores and reliable and representative measures of work performance. Good tests have all the data in their manuals about this issue. Often, work output/success is difficult to measure reliably and sensitively showing an individual’s personal contribution. It takes a very long time to get enough data to show that test scores do relate consistently and strongly to relevant work behaviour like productivity and satisfaction. Indeed, predictive validity, as it is technically called, is the criteria by which tests should be selected in the workplace.

5 Tests do not measure really important things like integrity and motivation.

This is simply not true as there are many tests of both integrity and motivation as well as



a myriad of concepts from entrepreneurship to narcissism. Once consultants realise that there is a demand for some sort of tests, they are willing to supply it, even though they quite often are not prepared to put in the time and effort to establish their validity. Once a “new” concept like agility or resilience becomes trendy, there will be those willing to measure it. Perhaps, we need a test on disruptive ability.

6 People change a lot over time.
Again, the data show the opposite. We know that people become a little less extraverted and neurotic as they get older, and a little more conscientious and agreeable; but the changes are relatively mild. After the mid-twenties, there is surprisingly little change in personality to the mid-seventies. Occasionally, some people experience significant trauma which does change them, but this is relatively rare. People feel they have changed a lot but the data says otherwise. For those in doubt, attend a school reunion.

7 All tests are much the same: none outperforms the other.
Tests trying to measure the same thing, like intelligence or personality can be radically different. Many share an approach and for some, traits or abilities look very much alike. Take for instance introversion-extraversion: many tests seem identical, but this trait can also be measured by weighing a person’s salivation after lemon is put on the tongue, or by what is called “the pursuit rotor test”. The question is not what the test looks like but rather evidence of their validity or reliability.

8 You can teach anybody to be a great performer.
The idea that anyone can become a brain surgeon or an airplane pilot with enough practice and training is still very popular. It is called the 10,000 rule and has been applied to athletes. But even the most radical of those who dismiss innate talent are forced to agree that you need a set of certain characteristics to succeed at certain jobs. You simply cannot teach any - or everyone to be good at serious technical and managerial jobs however hard they try if they do not have a particular ability and the desire to exploit it.

It is also important to note that job selection is essentially an arms race. For every improvement on the employer side, there will be a reactive step-up on the applicant side.

9 All tests are biased particularly with regard to sex and race.
Some, but surprisingly few, tests do show sex and race differences. There are more differences on ability tests than personality tests usually. Just because there may be some sex or race or age or culture differences, does not mean they are invalid but rather, they need to be used in a very particular way checking against population norms. Any bias occurs in how they are used, not in what they measure. In many countries test users are obliged to show data on age, race and sex differences and what they do if this occurs.

10 Tests do not spot “difficult people” well enough.
There are numerous “clinical” tests that set out to do just this. There must be 20-30 very well-established tests that measure “dark-side” variables. The data suggest that many “problem people” at work are the result, not of poor selection or of something very wrong with them, but rather the way in which they are managed. This is not to suggest that selectors should not look for evidence of pathology and dark side traits, but rather that they should not always blame the individual or the selector if people become “problematic”.

11 Attitude, knowledge and skill are more important than intelligence and personality at work.
Select attitude; train skill. If, by attitude, critics mean motivation then this is partly true. No matter how bright an individual or how well fitted they are to a job, if they are not sufficiently *intrinsically* motivated, very little can be done. Knowledge and skill can be taught: but this is affected by personality and intelligence. Brighter people learn faster. Certain personality types pick up skills faster than others. “Attitude” can be assessed and is very important.



12 The “old trio” (Application Form, Interviews and References) work well enough in selection.

Again, this is partly true if: the application form collects biodata that is important and relevant to the job; the interview is planned and structured; the references are collected from people who know the candidate and are prepared to tell the truth. These, however are rarely done and the old trio is woefully inadequate.

New Developments

Every generation of scientists has attempted to exploit the technology of their time. Researchers have noted the possibility of using the world wide web to do personality research. There are now companies who track huge numbers of people on the web and build various profiles, though not following any classical or modern personality theory.

The new assessment technologies (predominantly the web) have specific goals: improve efficiency, enable new screening tools, reduce costs, standardise the HR system, expand the applicant pool, promote the organisational image, and increase applicant convenience.

There are plenty of speculators and futurologists in this area, both academic and non-academic, the latter often being science journalists, practitioners and consultants. An example is McHenry (2017), himself both an academic and a test publisher made five assertions about the future of psychometric tests:

1. Smartphones will replace computers for employee assessment.
2. High-quality psychometric testing services will be sold directly to consumers.
3. Advances in the neuroscience of personality will reveal which are the most valid individual differences to measure and how best to measure them.

It is easy to be flooded with inappropriate applicants: people lacking the required and specified qualifications, experience or place of domicile apply online because it is so convenient, quick, and easy.

4. The digital badging movement, coupled with the use of big data and new forms of digital CV, will render many of the current applications for high-stakes testing redundant.
5. The basis for employee development will in the near future be derived from the data yielded by wearable devices and not from psychometric tests” (p. 268).

Ihsan and Furnham (2018) did a comprehensive review of five new technologies: big data, wearable technology, gamification, video-resumes and automated personality testing. They argued that at this stage, there is more *absence of evidence* of their psychometric properties of these new approaches, rather than evidence of absence of their validity.

However, there are also *unintended consequences* and effects on these developments. Thus, the use of the internet does expand the applicant pool but also increases the number of under-qualified and out-of-country applicants. It is easy to be flooded with inappropriate applicants: people lacking the required and specified qualifications, experience or place of domicile apply online because it is so convenient, quick, and easy. There is also the loss of personal touch that both assessor and assessee value and respect. There are still concerns about adverse impact, which means that certain groups simply do not have access to the technology to take the tests.

Chamorro-Premuzic et al. (2016) predicted that “profiling tools will become invisible to individuals and require no deliberate attention from job applicants or incumbents. Most people will be profiled already, and if they aren’t, assessment will operate in the form of covert or subtle algorithms embedded in other activities, including fun and interactive, game-like experiences” (p. 39). They have also further highlighted that there is little or no academic research for some of these methods, suggesting that the validity and reliability of these tools are still unestablished (Winsborough & Chamorro-Premuzic, 2016).

It is also important to note that job selection is essentially an arms race. For every improvement on the employer side, there will be a reactive step-up on the applicant side. Applicants learn how to clean their Facebook profiles and photos.



The history of psychology is littered with examples of how the primary technology of the time shapes not only theories and technology but also how many promises were never delivered.




Therefore, the question is whether it is widely known that your social media will be scraped for employability data. Many people will either create a dark social media presence or go off-the-grid for information that might be coded in a negative fashion.

Psychometricians know how expensive the validation process for any technique is. To establish various kinds of reliability (test-retest, alternative forms, internal) and validity (concurrent, construct, discriminant, incremental, predictive), it is necessary to collect considerable amounts of data (Furnham, 2008).

Finally, Ihsan and Furnham (2018) have noted that the new technologies present a range of new problems. The first is obtaining data about an individual without their consent. There are important questions about the accuracy of these data and who indeed is supplying this. Next, there are many potential issues around “wearables” and issues around surveillance. What if people refused to put on these wearables? How would an individual feel about an employer who has an electronic daily map of whom they were in contact with? Wearables can give physiological data as well as contact data. In this sense, they provide data on wellness and physical fitness, which may be considered an inappropriate and unethical way of assessing people. They argue that ethics committees need experts and specialists on these new technologies to inform other committee members. Also, the committees could benefit from the insight of users and researchers who themselves do not have any financial interest in the development or sale of those technologies.

It has been said of some psychological tests that they are techniques in search of a theory or solutions to non-existent problems. There is always a danger when exploiting the possibilities of new technology that sufficient evidence is collected and assessed to

show their incremental validity over existing and established methods. The history of psychology is littered with examples of how the primary technology of the time shapes not only theories and technology but also how many promises were never delivered. 

About the Author



Adrian Furnham is Professor at the Norwegian Business. He is also the Principal Psychologist at Stamford Associates in London. He has just finished his 93rd book

References

1. Chamorro-Premuzic, T., Winsborough, W., Sherman, R., & Hogan, R. (2016). New Talent Signals: Shiny New Objects or a Brave New World?. *Industrial and organizational psychology: Perspectives on science and practice*, 9, 621–640.
2. Furnham, A. (2018). The great divide: Academic vs Practitioner criteria for psychometric test choice. *Journal of Personality Assessment*, 100, 498-506.
3. Ihsan, Z., & Furnham, A. (2018). The new technologies in personality assessment: review. *Consulting Psychology Journal: Practice and Research*, 70(2), 147-166.
4. McHenry, R. (2017). The future of psychometric testing, In B, Cripps (Ed). *Psychometric testing: Critical perspectives* (pp 269-281). London; Wiley.

