



why assessment centre validity is falling

and what to do about it

“ In practice assessment centres have degenerated into tick-box assortments of off-the-shelf exercises...far from being the haute couture of the assessment world, they have become high street rip offs.”

Steve Blinkhorn

Overview

Like a high potential manager, the assessment centre method showed impressive promise at an early stage in its career. But like the derailing executive who assumes initial success will ensure future success and fails to deliver consistently against expectations, assessment centres seem to have lost their way.

At an early point the future looked bright for the assessment centre method. As well as providing significant predictive power, its career was assisted by the finding that assessment centres reported low adverse impact across different groups, an issue which cognitive aptitude tests had struggled with. Assessment centres were judged to be unbiased and fair, and candidates saw them as more relevant, realistic and face valid than the alternatives (e.g. biodata or psychometric tests). But even this career appeal is diminishing, as it isn't clear that assessment centres do in fact deliver lower adverse impact¹.

Assessment centres have in the past shown excellent power to improve the decisions we make about a key factor of organisational success: the leaders we recruit, promote and appoint to senior positions. But a combination of dynamics is undermining their current effectiveness.

Faced with leadership derailment, the wise executive coach Marshall Goldsmith suggests “what got you here won't get you there”. If assessment centres are going to get their career back on track, they need to rethink the dynamics of their success and failure and reposition themselves for a different future.



¹ “Ethnic and Gender Subgroup Differences in Assessment Centre Ratings: A Meta-Analysis” Dean, Journal of Applied Psychology, 2008

The emerging problem

Assessment Centres are an established part of many organisations' talent management strategy. Assessment centres are widely used as the selection method for the intake of new talent, the promotion mechanism for the career development of high potential managers and professionals, as well as the process for making senior level appointments.

The assessment centre method - the mix of different exercises simulating key aspects of management and leadership life, in which candidate performance is observed and evaluated by multiple assessors to produce a profile against different dimensions - has become the established gold standard of assessment.

Last year a meta analysis¹ to collate the research found that the validity of assessment centres, i.e. their power to predict future performance was falling. From a reasonable validity of .37, reported in Gaugler's 1987 meta analysis, Thornton and Gibbons' collation of the research in 2009 indicated that assessment centres are now reporting much lower average validity coefficients, more of the order of .27. Or put another way: assessment centre performance may explain less than 10% of the variation of current and future performance.

Given that a 30 minute paper and pencil test of cognitive aptitude² possesses predictive validity of around .5 (or 25% of the predicted variance in effectiveness) something strange is going on. What?

¹ "Validity of assessment centres for personnel selection", Thornton et al, Human Resource Management Review, 2009

² "The predictive validity of cognitive ability tests: A UK meta-analysis", Bertua et al, Journal of Occupational and Organizational Psychology, 2005

This article summarises the research base in assessment centre validity and also draws on the results of a short survey of the "wisdom of the crowd" of practitioners and experts to explore:

- what dynamics are influencing the power of assessment centres to improve the resourcing and development decisions we make?
- what are the practical implications for the future application of assessment centres?



Falling validity



“The reasons for the ostensible decline in validity over the past 40 years are not clear.”

Thornton & Gibbons, 2009

For some the finding of declining validity was no surprise.

Assessment centres have always been problematic. The “theory” that leadership effectiveness is the combination of competency dimensions, and that these criteria can be assessed across the different exercises of an assessment centre had never been established. Instead the opposite had been reported: greater consistency for different competency dimensions within the same exercise than across the same competency on different exercises¹.

This “problem” isn’t statistical nitpicking, it is fundamental to assessment centres and how they have been designed to operate. Rather than yielding meaningful measures of competency dimensions, assessment centres seem to measure performance on the challenges of specific exercises.

But as long as assessment centres worked, i.e. they predicted subsequent work performance, no one was too troubled by how they worked.

For many consultants, practitioners and organisational users Thornton and Gibbons’ 2009 meta analysis was unwelcome news. Not only do assessment centres not work in the way the theory suggested, they might not now be working very well at all.

¹ “Why Assessment Centers Do Not Work the Way They Are Supposed To” Charles Lance, Industrial and Organizational Psychology, 2008

The history of assessment centres

It was the German Wehrmacht who were the pioneers of what is now recognisable as an assessment centre (a three day event comprising different exercises and tests, observed by a multi-disciplinary panel), a development that was in turn picked up by the British military when they introduced the War Officer Selection Boards. Regarded as a successful innovation, the Civil Service Commission applied the method of assessment centres for non-military recruitment.

In the US the AT & T Management Progress studies in the 1950 - 60s became the trigger moment for the evolution of the assessment centre method. Originally designed as a research programme to track the career development of new recruits and identify the indicators of rapid progression, the AT & T exercise also became an experiment in assessment centres' capacity to predict the future.

Because the assessment centre results for the population were kept confidential, amazingly to individuals and their line managers, there was no risk of the "self fulfilling prophecy" confounding the research (those predicted to progress do in fact progress faster but only because the predictions influence progression decisions).

An eight year follow up indicated a correlation of .51 between overall assessment ratings and subsequent job performance ratings¹.

Organisations, excited by a breakthrough in new assessment methodology, were quick to exploit the potential of the assessment centre method within corporate life. Throughout the 70s and 80s, assessment centres were applied to a range of resourcing and development processes, from initial recruitment and promotion to fast track development.



¹ "Managerial lives in transition: Advancing age and changing times" Howard, & Bray, 1988

A survey of organisational practice

This article combines a summary of the desk top literature with a survey of organisational experience to identify the dynamics at work. Here we drew on the “wisdom of the crowd” of over 100 practitioners and users to put assessment centres through their own centre to evaluate their effectiveness and impact. In our survey we presented three “assessment exercises”:

1. what is the **candidate experience**? Not a definitive test, but an important insight nonetheless into the positioning of assessment centres, and how they are perceived by end user participants.
2. what is the **candidate cost**? Assessment centres have been recognised as an expensive assessment option, but the argument has been that the long-term gains out-weigh the short-term costs. Do we know how much they cost to run?
3. what is the **validity of assessment centres**? This is the “show stopping” exercise of the event. Do assessment centres actually meet the expectations of their claims? Or, is the assessment centre a method that once worked, now losing relevance to the challenges of identifying leadership potential in 2010? If so, why?

What do these results suggest for the application of assessment centres: a fundamental rethink or their total abandonment?

What are the implications for future progressive practice to ensure organisations are equipped to identify leadership potential and make informed leadership appointments?



Exercise 1: the candidate experience

Historically the research has indicated that candidates have perceived assessment centres as fair and relevant. They might not like them. Instead they seem to view them as a necessary organisational evil; but they do view them more positively than other selection activities.

Judging from the pattern of results from our survey of experts, this still seems the pattern in 2010.

Most participants find the process insightful and accurate, although more than one survey contributor noted there may be a skew in candidate responses post an assessment centre. Few candidates, particularly in a selection scenario, are likely to report highly critical views.

Candidate feedback is more positive when:

- there is an appropriate balance of support and challenge within the event
- the event provides an opportunity for feedback
- senior managers are involved in the process
- the exercises mirror real-life organisational challenges. Candidates dislike contrived or artificial simulations

Positioning is key to manage candidate expectations. In recruitment situations candidate work out the rules of the game: succeed and pass to the next stage. For internal career development programmes, there is greater scope for mixed messages and unrealistic expectations.

There remains scepticism about psychometric tests. Although measures of personality and cognitive aptitude do seem to enhance the overall predictive power of assessment centres, candidates remain unconvinced.



Exercise 2: candidate cost

Assessment centres, positioned as the gold standard in assessment, have always been an expensive option. Advocates however point to the improvement in overall utility, and that short-term costs need to be viewed against the longer-term benefits. This is right. Where the difference between excellent and average levels of leadership is translated into substantial variation in business performance, the cost of any valid assessment method is outweighed easily by the organisational gains.

How much does it cost to design and implement an assessment centre? We presented to our end user group of experts the scenario of a middle management programme for 100 candidates, attending a one and a half day residential event in cohorts of twelve.

The survey of assessment centre users generated a wide spectrum of responses. The more miserly of the survey group suggested a candidate cost of £130, and the more extravagant, £7500.

Removing the extreme outliers, most responses gravitated towards the £2 - 3000 mark. Because a number of the survey contributors outlined their cost assumptions in design, event management, candidate and assessor time, and follow through, we mapped out the different cost elements. £3000 per candidate seems a realistic estimate for a professional process.

A development programme for the full population of 100 managers would therefore cost £300,000.

Whether this cost is a shrewd investment or not depends on a number of factors, not least the validity of the assessment centre, our next exercise.



Exercise 3: validity

For exercise three in our assessment centre of assessment centres we asked the survey group:

- what validation have you conducted to track the relationship between assessment centres you've designed and implemented with subsequent work performance?
- what is your estimate of the validity of the assessment centre method, with the assumption of a professionally designed event that is fit for purpose? Where the correlation between assessment centre ratings and future leadership performance ranges from zero (no relationship) to 1 (a perfect prediction) , what is the view from the "wisdom of the crowd"?

A number of users within the survey group reported conducting specific follow up studies, some of which were impressive in their diligence to track progress over time. However this was very much a minority.

The "wisdom of the crowd" generated a spectrum of estimates of the validity of assessment centres, including outliers, one who suggested a perfect correlation to the more challenging suggestion of zero validity. In fairness to the challenging outlier, this practitioner had conducted an extensive validation exercise in one of the UK's largest employers and found no relationship between assessment centre scores and subsequent work performance.

The most frequent response was however a validity coefficient of .6, a confidence that assessment centres can predict 36% of the variance in performance effectiveness.

If Thornton & Gibbons' summary of the validity research is accurate, our survey group is much more optimistic than the evidence indicates.

So what factors are at work in explaining the "validity problem"?



The theory of declining standards

This is the straightforward explanation that we shouldn't expect assessment centres to maintain high levels of predictive power if they are being developed and introduced in a slap-dash manner. When short-cuts are taken in exercise construction, assessor training and event management, the evaluation of candidates can be nothing other than inaccurate.

It is true there is no shortage of embarrassingly bad examples of assessment centres, based on flawed leadership models and crude exercises observed by poorly trained assessors. But it is also true that the last two decades or so has seen a massive growth in professional activity, with hundreds of consultancy firms providing specialist services.

It is also unlikely that poorly designed and executed assessment centres would be incorporated in any analysis of their predictive validity. Sloppy practice in development is unlikely to translate into robust follow-up programme of evaluation. Indeed, it is more likely that any skew in sampling would be towards the inclusion of assessment centres based on higher standards of design and implementation.

Whilst there is reason to believe that sharp practice and expediency have undermined professional standards in assessment centres, this argument, reassuring to those who want to preserve the status quo, doesn't seem to explain the fundamental problem.



The theory that leadership life has changed

This is the argument that ratings made 10 years ago captured information about candidate performance that is no longer relevant to a different leadership world. Here the logic runs that the candidate skills and styles that made for high performance on assessment centres a generation ago have little relevance to a new set of demands and challenges. We shouldn't therefore expect assessment centre ratings to predict a different set of leadership criteria.

If this is true then we need to rethink the application of assessment centres. Bray and Campbell in the AT & T studies of the 1960s anticipated that assessment centre ratings would be a key predictor of future performance. It looks like we should be more modest in our expectations of what assessment centres can deliver.

A variation of this argument is that assessment centres encourage the kind of conformity that has squeezed out diversity of leadership outlook. When a senior executive¹ at BMW comments that the current generation of managers are “smart, competent and very nice...but almost totally interchangeable” and lack flair and vision, he is highlighting a worrying trend in how organisations identify leadership.

In the same article, researcher Sandra Siebenhüter puts the blame on assessment centres. “The problem is that centres measure only what is easily quantified. It may be that those who do well are simply well prepared or have the right stuff for the assessment centre and not necessarily for the job.” She calls this conformity “gleichmacherei”: “forcing things to be the same”. Assessment centres, designed to extend the pool of available talent, are now doing the opposite: generating “managers with no profile”.

These arguments have some merit.

The leadership world is different now, and it is entirely plausible that assessment centres have not been sufficiently sensitive to changes in leadership requirement.

And there is mileage in the suggestion that, in the spirit of assessment centre efficiency, the qualities of individuality and creativity have not been recognised fully. In the search for consistency and objectivity of measurement, assessment centres do set an expectation of conformity against a check-list of “approved behaviours”. If this is the case we shouldn't expect candidates with “no profile” to go on to be successful in a complex and diverse world that requires creative solutions.

However this theory is unlikely to explain completely how a method that delivered validities of around .5 in the late 1960s, .37 in the late '80s, is only yielding predictive power of .27 in 2009. The leadership world has changed, but that much?



¹ <http://www.telegraph.co.uk/money/main.jhtml?xml=/money/2008/01/16/cmmanage16.xml>

The theory of an out-dated model

This line of argument builds on the previous theory. Assessment centres assume that leadership can be summarised into a set of competency dimensions, and exercises constructed to measure those dimensions can map out a spectrum of effectiveness in the evaluation of overall candidate performance.

Of course key capabilities matter and can and should be assessed. And a professionally designed and implemented assessment centres will provide an important overview into these capabilities. But if the latest evidence demonstrates that they can at best pick up only 10% of the variation in effectiveness, it's clear they're missing something important.

What they miss is:

- a sense that **leadership success is contextual**¹. The days of “martini management” are over. The leader who can take on any challenge, “any time, any place, anywhere” is rare. Instead there is a focus on leadership which excels within the complex dynamics of specific challenges. The kind of executive who relishes the opportunity to expand the business quickly is a very different individual to the leader who can trouble-shoot problems and turn around a failing business unit. The typical attempt by assessment centres to reduce the richness of the specific assessment data into an Overall Assessment Rating misses the point: there is no all-singing-all-dancing leader who will be equally competent in taking on any leadership challenge.

- the **importance of wisdom**² as an important component of leadership effectiveness. Wisdom won't be assessed in a 20 minute situational judgemental task. Wisdom is that theme of leadership life that manages complexity, uncertainty and ambiguity, uses judgement to trade off the pros and cons of different options, balances what is ideal with what is possible, and copes with the pragmatics of organisational life with emotional maturity. It seems difficult to see how the standard assessment centre that speeds candidates from exercise to exercise with the expectation of immediate impact can do justice to a quality of leadership that is becoming increasingly critical to organisational success.

This seems a compelling set of arguments.

Assessment centres have assumed leadership comprises a set of competency constructs, and the combination of these constructs adds up to a score that will predict future overall effectiveness. But since the early days of assessment centres, we have discovered that leadership is more complex, that it is played out within the dynamics of different kinds of challenges.

Instead of looking to identify the candidate with the highest set of overall scores we should identify the specific leadership patterns which will excel in specific leadership situations.

¹ <http://www.azureconsulting.com/files/1/24127127/whatsOnTheRadarScreen.pdf>

² <http://www.azureconsulting.com/files/1/91276339/SevenPillarsofWisdom.pdf>

The theory of candidate tactics in impression management

We have “How to Succeed at an Assessment Centre: Test Taking Advice from the Experts” and “Succeeding at Assessment Centres for Dummies” to map out the tricks of the trade and help highly motivated and ambitious candidates put their best foot forward in assessment centres.

There is also an array of web sites providing general advice as well as blogs analysing the detail of specific organisations’ assessment centre practices. One candidate reporting his views of Accenture’s process suggests the following strategy: *“as long as you don’t say or do something stupid, it’s going to be a case of how much they like you, and more importantly how you’d fit within their corporate frame. If you’re extremely intelligent or have a lot of personality, tone it down.”*

If assessment centres are being skewed by shrewd impression management, we shouldn’t be too surprised that a process which once worked in the relative naiveté of corporate life in the 1960s isn’t cutting the mustard in the knowing world of on line social networking in the 21st century.



Of course the stratagems of career advancement are part of any leadership repertoire. We know from our research into Career Tactics¹ that smart tactics make a difference. Any aspiring leader works out the “rules of the game”, recognises the players and the importance of stakeholder management and political influence and adapts their approach accordingly. The argument here isn’t that tactics don’t matter in leadership life. They do. But they shouldn’t matter that much. And when they do, self-seeking advancement drives out authentic leadership that looks to make a genuine difference.

The proposal that assessment centre validity is declining because some candidates “out perform” their less savvy but more gifted candidates has considerable merit. Even worse, candidates of genuine talent, unimpressed by the antics of their self-seeking peers looking to manipulate the game, switch off from the entire activity.

And support for this argument comes from the evidence that the last 10 - 20 years (an era in which assessment centres were firmly established in selection and promotion decision making), has not so much witnessed a leadership renaissance but seen a decline in leadership credibility and levels of employee trust.

¹<http://www.amazureconsulting.com/files/1/69156045/CareerTacticsLeadershipEffectivenessProgression.pdf>

Implications

Each of these theories advance a plausible argument. And it is probably the interplay of different factors that is the best explanation of the decline in validity.

No doubt there are excellent assessment centres - based on a forward thinking model of future requirements, that provide the time and space for the qualities of creativity and wisdom to be displayed, encourage candidate diversity and minimise the impact of tactics in impression management - working well to improve decision making for the leadership future. But this doesn't seem to be the overall pattern.

Faced with this finding, what are the options?

Abandon. This has been the strategy of a number of organisations in our survey. Reviewing the costs and benefits they have concluded that the assessment centre is unwieldy and failing to provide a sufficient return on investment. Jettisoning an expensive practice that isn't working is obviously sensible. Abandoning the assessment centre and replacing it with less valid alternatives is not¹.

Fine tune the existing model. A number of survey contributors outlined specific ways to improve design and implementation. Here the argument is that, if predictive validity has declined, we need greater professionalism of practice and rigour in the construction of the activity. And if we can incorporate innovation in technology (e.g. virtual exercises) so much the better.

Rethink and reposition. This seems the most sensible course of action. The assessment centre method requires more than a fine tune to grapple with some fundamental questions:

- what can and can't be measured?
- when do assessment centres add genuine value and when do alternative methods provide a more cost-effective option?
- should assessment centres focus more on exercise performance rather than attempt to profile competency dimensions?
- how should assessment data be captured, consolidated and integrated with other information to optimise decision making?



¹ "Utility of the assessment centre as a selection device", Cascio & Silbey, Journal of Applied Psychology, 1979

Improving assessment centre validity: 5 suggestions

1. Don't over elaborate job analysis or competency profiling

To the purist practitioner, this is heretical. The classic recommendation is to conduct a comprehensive analysis to map out a competency framework that becomes the criteria for assessment centre design. And for assessment centres used in recruitment for specific roles of course the process needs to reflect the key criteria for success within the role.

But for assessment centres being used as part of talent identification within a leadership development programme, instead of expending huge amounts of organisational time and energy to generate a competency profile which will look pretty much like any other organisation's and probably change in two to three years' time, it might be better to structure the assessment around four key questions¹ :

- is this individual **credible**?
- does the individual display exceptional **capability** in any specific areas?
- does the individual demonstrate the values, resilience and courage that underpin leadership **character**?
- is the individual proactive in their **career management**?



¹ "Rethinking Leadership Realities":
<http://www.amazureconsulting.com/files/1/73424767/RethinkingLeadershipRealities.pdf>

Improving assessment centre validity: 5 suggestions

2. Measure what can be measured and measure it well

Assessment centres are typically based on a design of 10 or 12 competency dimensions and involve the construction of exercises to differentiate the nuances of these distinct dimensions. But it seems we may have been wasting our time.

Despite scores of studies, not one finding has indicated that assessment centres in fact measure more than performance in three or four areas. Factor analysis is that statistical technique that, rather than being impressed by the theory of our leadership model, examines if it holds up in practice. Specifically it crunches the numbers from the assessor ratings to ask: can assessors provide meaningful evaluations that replicate the proposed leadership model?

In no research study has the answer ever been yes. Indeed in one analysis¹ assessor ratings on four dimensions had a higher correlation with job performance than the correlations based on more dimensions. “Less is more” in assessment centre design and implementation.

Complex assessment centres may provide the emotional reassurance of rigour. The reality is that simplicity focused around 3 - 4 highly relevant exercises may provide more predictive power.



¹ “A meta-analysis of the criterion-related validity of assessment center dimensions”, Personnel Psychology, Arthur et al , 2003

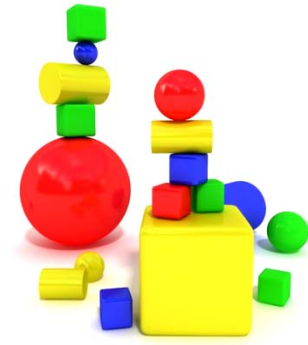
Improving assessment centre validity: 5 suggestions

3. Only use exercises that allow genuine diversity and creativity of response

The requirement for consistency and objectivity (under time pressure) of evaluation set the tone for a generation of “tick box” assessment centres. Here evidence is recorded against a prescribed exercise score sheet and set of competency indicators.

Convenient for assessor reliability and efficient for process, this kind of assessment event is unlikely to provide an important insight into candidate effectiveness in taking on complex and ambiguous challenges.

This isn't to suggest that exercises - group discussions, fact finding problems, role plays and so on - need in themselves be complicated. Indeed, the best exercises ask a straightforward but tough question. It is to recommend however that more time is allocated to the design and selection of high impact, relevant and challenging exercises than the crafting of competency criteria.



Improving assessment centre validity: 5 suggestions

4. Improve the information flow of data capture, consolidation and integration with other information

Practitioners have been proud of the principle that assessment centres create a “level playing field” in which all candidates have the same opportunity to demonstrate their talents and skills. And it is right and proper that assessor views of candidate performance are not biased by any previous views of effectiveness.

But the consistency of assessment centres shouldn't dismiss all other information as likely to contaminate the objectivity of the assessment. Meaningful recommendations - whether it is for assessment or development - need to put assessment centre ratings into context, the context of career achievements, work performance outcomes, 360 feedback data and psychometric test results.

We now have a decent knowledge base of the pros and cons of assessment centres, 360 feedback and psychometric data, and the incremental validity that is achieved through an informed integration of the different inputs. The gain comes from using this knowledge to improve the decision making - either in selection or promotion decisions or in guiding the development of high potential candidates.



Improving assessment centre validity: 5 suggestions

5. Get rid of the final assessor conference

The conventional approach is the “assessor wash up”, that forum in which all assessors meet to review the findings and agree competency scores and achieve a consensus over overall assessment ratings. Alternatively, the results from the different assessment exercises are consolidated using a series of algorithms to weight and summarise the exercise-criteria ratings.

In the absence of any evidence that “wash up conferences improve the predictive power of the assessment centre method, and given the time and associated costs of these events” it may be better to integrate the data statistically.

Cancelling the assessor wash up won't undermine validity (it might optimise it), but by reducing costs, it will certainly improve overall utility.

“The simple mechanical combination of assessor data apparently does not reduce the predictive power of the assessment centre method.”

John Bernadin



About us

AM Azure Consulting Ltd works with a broad portfolio of clients in the design and implementation of on line services in recruitment and selection; management assessment, development and career management; on line leadership tool kits, 360° feedback, performance management; and talent and succession management.

If you are interested in our approach to leadership assessment and development, our assessment tools and our on line resource at www.theleadershipdirectory.com call us:

- 44 (0) 1608 654007 or email
- admin@amazureconsulting.com

- **we're professionals but we're not pompous.** We are at the edge of the latest research and thinking in the field of people management, but we're not precious about the "one thing". We have some good ideas to help your organisation perform even better, but we know that you have some better ones, but want support in making them work. We don't impose the "solution".
- **we design but we also implement.** Our content, design and technology can build cost effective solutions quickly. Our consultancy experience of "real world" implementation and our levels of client service will move things forward from initial concept to results rapidly.
- **we start things to build momentum but we also follow through.** Results come from the discipline of "making it stick", of evaluation, learning and continual improvement. And we maintain ongoing relationships with our clients to keep achieving positive outcomes.

