

# The A&DC Thought Leadership Series

## Assessor Discussions

Are Wash-Ups on the Way Out?



## **Assessor Discussions**

### Are Wash-ups on the Way Out?

#### **Introduction**

The “wash-up” or assessors’ discussion, in which assessors come together at the end of an assessment centre to discuss their observations and arrive at a final set of scores, has been an integral part of assessment centres for many years. However, some researchers are now starting to question the value of this process and advocating that it should be done away with. This article looks at the evidence from a practitioner’s point of view and argues that wash-ups still have a key role to play in ensuring the accuracy, consistency and fairness of assessment centres.

#### **Recent Concerns about wash-ups**

In a study presented at the British Psychological Society’s Conference in January 2006, Dr Chris Dewberry of Birkbeck College and Deborah Jordan of Ernst and Young described a series of wash-ups they had observed (they use the term “consensus meetings”) and expressed a number of concerns. They report evidence of “latent-informal processes” which had nothing to do with the design or the intended content of the assessment centre. In particular, they were concerned by three things. These were that, in the course of wash-ups:

- 1) “Assessors make active attempts to persuade others that a candidate should, or should not be selected.”
- 2) That these attempts to persuade are often based on “general impressions of candidates” (rather than on evidence of the agreed competencies) and
- 3) That assessors who are more senior use their seniority to persuade other assessors to accept or reject candidates.

Using a combination of the data from their observations and statistical analysis of the scores from the centre as well as the first interviews used, they concluded that bearing in mind: “the significant costs associated with running consensus meetings; there seems little or no justification for their continued use.”

#### **Observations on Dewberry and Jordan’s study**

This seems like a far-reaching conclusion for a variety of reasons. Firstly, their study appears to be based on observations of only four wash-ups in three organisations and, although they report having noted consistent trends, the qualitative evidence quoted is only based on one organisation’s wash-up process. Admittedly, this was a large and successful multinational organisation, which we are told recruited about 400 candidates a year. However, just because an organisation is “large and successful” and recruits large numbers it does not necessarily follow that it observes best practice in the design and delivery of assessment centres.

The example described seems to have a number of serious flaws in the process used for the wash-up, and these could all have been addressed, as I will describe later.

In addition, graduate recruitment is usually the largest-scale recruitment that organisations conduct. As a result, cost is an important consideration and often the most junior HR staff are encouraged to “cut their teeth” on graduate assessment centres. There is therefore a real danger that inappropriate evidence or evidence of a “seniority” effect could appear.

However, despite the claim that this organisation had “proper assessor training and good professional practice in the running of ACs” there appears to have been no agreed process for the wash-up, with assessors saying wildly different amounts about each candidate. One simply reads out their scores (with no attempt to provide supporting evidence) while another launches into a lengthy and rambling discourse, which contains all manner of anecdotal evidence and subjective opinions. One wonders, at this point, what the facilitator or chair of the wash-up was doing.

In another, almost farcical exchange, two assessors appear to engage in horse-trading.

One had awarded a two (below the pass-mark) for communication skills in a group exercise and appeared to have evidence to back this up (they say that: “she did really withdraw from the meeting...it was quite a large period when she didn’t actually say anything at all....”). By contrast, another (more senior) assessor had been extremely impressed by the same candidate in an interview and had awarded a 4.

One of the key principles of assessment centres is that the exercises are a discrete series of observations, so these two scores could quite reasonably co-exist. Indeed, the more junior assessor points this out, suggesting that: “that might demonstrate that she actually she is more comfortable in a group situation.”

However, rather than seeking to establish the accuracy of each score separately the facilitator asks the more senior manager: “Would you be happy to go down to a 3?” and, when he refuses, the more junior assessor suggests: “I’ll happily change mine to a 3 or more.” This is clearly a case of influence being exerted by more senior manager. However, once again, one wonders why the chair of the wash-up did not seek to remind the two assessors of the key principle that the scores from each exercise should be viewed in isolation until they have been agreed and that only then should an overall score be considered.

I will return to this example later and discuss some ways in which the problems described can be avoided by those wishing to tighten up on their assessment processes. For the moment, however, I would like to look at what “added value” (if any) wash-ups add to any selection process.

### **Why bother having a wash-up?**

Having been involved in designing and running assessment and development centres for a large number of public and private sector organisations over the last eight years, I have come across some compelling reasons for continuing to use wash-ups.

## Differences in Ability and Motivation between Assessors

One of the most important reasons is that assessment is a skill which managers within organisations have different degrees of aptitude and enthusiasm for. Many managers become involved in assessment centres because they care passionately about recruiting the right talent into their organisation as a means of securing its future. As a result, they are highly conscientious and thoroughly professional. Unfortunately, not all managers who put themselves forward as assessors fall into that category. Other managers have given a wide variety of reasons for putting themselves forward. These include:

- That they have been ordered to do so
- They enjoy it
- They see it as “something good to have on your cv” and therefore as a means of getting promotion and
- Because it “gets them out of the office” often to another location and “away from the day-job” for a short break.

Given this wide range of reasons for assessing, it is logical that non-professional assessors will approach the process in different ways and this can lead to inconsistency of scoring.

## Causes of Inconsistent Scoring

I have attended and chaired a large number of wash-ups and have frequently observed inconsistency of scoring among assessors from the same organisation. This can occur for a variety of reasons but some of the most widespread problems include:

- **Errors in classification** (whereby an assessor has wrongly categorised a piece of behavioural evidence so it is considered as evidence of another competency). For instance, evidence of Communication may be included under Leadership. Any uncertainty over the definitions and boundaries of the competencies can be addressed in the wash-up.
- **Not fully understanding the content of an exercise or the task set.** One recent example I came across was an analysis exercise in which candidates were asked to study a large amount of data and then make some recommendations on possible courses of action for the organisation described. Neither the candidate's report nor their subsequent presentation contained any recommendations at all yet the assessor (a senior HR Manager) awarded scores at the “effective” level.
- **Subjective data**, such as interpretations of body language or attempts to ‘read candidates’ minds’ are sometimes included e.g. “she obviously lost interest at that point.”
- **Misunderstandings of the required standards.** For instance, assessors will often award a good score on the grounds that a candidate was “the best one we saw today” rather than because they meet the required standard for a target job. This can work both ways, of course, with assessors sometimes having unreasonably high expectations that go beyond what is required in a target job.
- **Different standards of note-taking.** Assessors can also vary greatly in their willingness to take notes or to record behavioural examples. Some hardly see the point and the wash-up plays an important role in insisting that they supply specific evidence to justify the scores they have awarded.

- **Limited time available for Assessor Training.** One of the studies cited by Dewberry and Jordan is the work by Caldwell, Thornton and Gray (2003) which describes 10 things that can cause problems with assessment centres. One of these factors is assessor training and Dewberry and Jordan refer to Spychalski et al's (1997) study in the US in which assessor training was found to last "about four days on average." They therefore conclude that weaknesses in assessor training cannot be behind the lower than expected predictive validity of many assessment centres.

However, this seems very far removed from reality in the UK. In my experience, busy managers are reluctant to spend more than a day or two away from their desks and in some financial organisations, where time literally is money, 2-3 hours may be all that they can offer. With such big variations in the time available, it is hardly surprising that assessors' levels of skill and knowledge will vary. However hard they try, it is very difficult for them to fully understand a set of exercises and score them consistently after only a few hours of training.

- **Benchmarking.** One of the most important contributions made by a wash-up is to provide a forum for agreeing standards or benchmarking. When a group of line managers get together for the wash-up, they jointly develop a view of how much of each competency they would like to see from candidates to be confident that they will be successful in a particular role. This benchmark then helps them to avoid the danger that there are harsh or more lenient assessors among the group. Consistency among the assessors can be increased when they are in a position to say: "well, we gave that candidate a 3 for doing x, so this candidate (who has done the same thing) should get the same score." Without the assessors having the opportunity to hear what other candidates have done in the same exercises, this process cannot happen.

For all of these reasons, the assumption that all of the assessors on an AC will score consistently and accurately seems like a major risk, and the prospect of running an assessment centre without a wash-up concerns me.

### What can be done to improve the quality of wash-ups?

There are therefore some major advantages to including a wash-up as part of an assessment centre. Yet Dewberry and Jordan highlight a number of potentially significant problems. The question is, what can be done in practical terms to avoid some of the pitfalls they describe?

- **Make assessor training a pass/fail event.**

Having run many assessor training events, it is often apparent that some managers are unable or unwilling to do the job to the required standards. For instance, they may refuse to keep adequate notes, they may write down their subjective impressions rather than verbatim quotes, they may be unwilling to include any behavioural examples as part of their write-ups, they may make sweeping generalisations or simply misunderstand the competencies and so on.

The trainer can often tell that at this stage that these individuals will undermine the objectivity of an assessment centre. Nonetheless in most cases they are allowed to continue as assessors, causing potentially serious problems for centre managers and fellow assessors, as well as undermining the quality of the assessment centre itself.

The ideal answer is to make assessor training a "pass or fail" event so that those unwilling to commit to following the agreed procedures can be "weeded out."

However, the reality is that many organisations are so short of volunteers to act as assessors that they dare not turn anyone away. Unfortunately, this leads to a great deal of stress and disruption on assessment centres as colleagues realise that they are being asked to work with someone who is not recording and scoring in the same way as everybody else. The “pass / fail” approach to assessor training would go at least some way to eliminating the risk of assessors relying on their status and “general impressions” as the basis of their scores.

- **Have a properly-trained Facilitator to chair the wash-up**

Another way of mitigating some of the risks highlighted by Dewberry and Jordan is to have a trained facilitator or chair for the wash-up. In the wash-up described in their study, active persuasion, general impressions and rank are all allowed to come into play. They describe the event as “a well-run assessment centre” but seem to overlook the fact that if they had had a chair or centre manager who was experienced and assertive enough, they would have stepped in to remind the assessors that much of their behaviour was unacceptable and contrary to the principles of an assessment centre.

For instance, when general impressions were being provided (one assessor said “I thought she was excellent...one of the best candidates I’ve interviewed in a long time”, and “I thought she was great”) the chair could have said: “I’m sorry, but let’s stick to what evidence we have of the competencies we are setting out to measure.” Similarly, in response to the senior manager exerting pressure on another assessor to change his/her scores, the chair could have pointed out: “well he/she observed the candidate in the group discussion and you didn’t!”

- **Agree on procedures for the wash-up (and stick to them!)**

In another part of the study where we are told that one assessor was out of the room and therefore the wash-up of a particular candidate could not be completed. The chair is reported to have attempted to move on to another candidate, recognising that no decision could be taken without all of the data but one of the assessors repeatedly says: “Definitely an offer, definitely an offer.” Instead of telling this assessor (and the rest of the group) that they must follow the same process for each candidate and that it is inappropriate to reach any such conclusion until all of the evidence has been heard, the facilitator climbs down and actually joins in, saying: “based on what you two have said it looks like it will be an offer, yeah.” While Dewberry and Jordan point to this as “a graphic example of active persuasion by an assessor” it seems to be an equally graphic example of a chair who makes no attempt to insist upon proper procedures.

- **Be Realistic about what you can measure**

It also seems that the design of the event contributed to some of the problems highlighted. The assessment centre consisted of only three exercises, yet it set out to measure 8 competencies. This is an ambitious goal, since the BPS guidelines favour a maximum of 4-6 competencies per exercise and also advocate measuring each competency at least three times.

- **Be aware of the problems associated with combining data from simulations and interviews**

The design of the assessment centre described was made up of a group discussion, a case study and a competency-based interview. The wash-up appears to have been dominated by the evidence from the interview, with the interviewer introducing a large amount of subjective information which had little to do with the competencies.

It is always difficult to combine the evidence from an interview with behavioural evidence from simulations as often anecdotal evidence creeps into the wash-up and can distort the other

assessors' view of a candidate. For instance, one client I have worked with had an assessment centre which initially included four simulations plus an interview. This set out to measure, among other things, the ability to cope with change. The organisation wanted to predict the applicants' ability to adapt to change as the senior managers of the future but, since the applicants were graduates, they had few examples outside of their education and family lives. As a result, many told harrowing stories of how they had coped with experiences such as bereavement or their parents' divorce, and this led some assessors to feel sorry for them and to want to improve their marks in other exercises as if to compensate in some way. This client soon decided that the interview data was having too much of an impact on the wash-ups and has now decided to concentrate solely on using simulations.

- **Consider each competency in isolation**

Dewberry and Jordan also describe how all of the scores for each candidate were: "written up on a board, observable to all by the facilitator." This practice also seems likely to lead to the assessors forming what the researchers call a "general impression" of each candidate before the wash-up started. This is easily avoided by the facilitator writing up the scores before the wash-up and then revealing only one line at a time as each competency is discussed.

- **Use Assessors of the same or similar levels**

The issue of differences in seniority causing more junior assessors to defer to their more senior colleagues also seems avoidable. To remove the risk of this happening, assessors should be chosen who are at the same or similar levels within the organisation.

Ideally, they should be one or two levels above the level of the target job so that they understand the role fully and also how it fits into the organisation.

## **Conclusion**

There is no doubt that wash-ups are "short-lived social groups" as Dewberry and Jordan describe. As such, there is an inevitable risk that individual assessors may try to influence one another and that social factors, such as degrees of power, may come into play.

However, these problems are not insuperable, and many of them can be dealt with swiftly provided an adequately trained chair or facilitator is involved. Many of the other risks described (e.g. the dangers of having assessor with differing degrees of seniority) are also easily avoided.

Bearing in mind the many ways in which inconsistencies of scoring between assessors can arise, the prospect of ACs without wash-ups is one that should concern those involved in running assessment centres.

Based on my experience, I am not at all convinced by the argument that there is "little or no justification for their continued use." Wash-ups serve an important purpose by ironing out inconsistencies and misunderstandings and by providing the opportunity for assessors to benchmark and agree on consistent standards.

## About A&DC

Founded in 1988 by Nigel Povah, author of 'Assessment and Development Centres', A&DC is one of the leading experts in the Assessment and Development field. We combine our expertise in business psychology and behavioural change to create and deliver end-to-end solutions to HR issues across the talent management spectrum.

We work in partnership with our clients to unlock human potential, using best of breed processes to enable them to select, promote and develop talented people who can contribute effectively to business growth and cultural enhancement. Always, we apply recognised best practice, putting our clients in a position where they can minimise risk and optimise return on investment in these critical areas of people strategy.

Based in Surrey, our Consultants operate across the UK. Through our international partners, we ensure that our comprehensive portfolio of products and services is delivered through specialists with a high degree of local cultural and business insight.

Assessment & Development Consultants Ltd

3 Lammas Gate, Meadow, Godalming, Surrey GU7 3HT UK

Telephone: +44 (0) 845 6888782 Email: [info@adc.uk.com](mailto:info@adc.uk.com) Web: [www.adc.uk.com](http://www.adc.uk.com)

A&DC® is a trademark of The A&DC Group Limited in the UK, USA and/or other countries around the world. All rights reserved.

