Development centres - a review of assumptions
Peter Carrick; Richard Williams
Human Resource Management Journal; 1999; 9, 2; ABI/INFORM Global
pg. 77

Numerous accounts (Boyle et al, 1993; Industrial Society, 1996) indicate the popularity of the assessment centre (AC) as a tool for development-related purposes. The popularity of what have become known as development centres (DCs) seems to stem, at least in part, from the demonstrated criterion-related validity of the AC method and much has been claimed for DCs in terms of the benefits they offer. However, little evidence exists to support the claims made. It would seem that practice in this area is heavily dependent on several assumptions. Our purpose is to analyse those assumptions in order to help practitioners identify areas of difficulty and to better understand what realistically can be expected of DCs. In doing so we draw on different bodies of literature and seek to bring together writings on ACs/DCs with those on aspects of career and employee development.

Development centres have their origins in AC practice. Although there is probably no such thing as a typical AC (or DC), we do have some standards which put forward characteristics that ACs should possess if they are to be worthy of the name. Thus, the Guidelines and Ethical Considerations for Assessment Center Operations, developed and endorsed by practitioners and academics in the US, define an AC in the following terms:

An assessment centre consists of a standardised evaluation of behaviour based on multiple inputs. Multiple trained observers and techniques are used. Judgements about behaviour are made, in major part, from specifically developed assessment simulations. These judgements are pooled in a meeting among the assessors or by a statistical integration process. In an integration discussion, comprehensive accounts of behaviour, and often ratings of it, are pooled. The discussion results in evaluations of the performance of the assesses on the dimensions or other variables which the assessment centre is designed to measure.

Although the US guidelines are expressed in terms of AC operations, they apply no less to DCs as they recognise that one application may be 'diagnosis for development'. This leads to the question of purpose. ACs, particularly in their earlier manifestations, concentrated on the assessment or measurement of individuals for such employment decisions as selection or promotion. These remain important purposes today. For example, Boyle et al (1993) report that slightly over half of the 402 large organisations (over 1,000 employees) responding to their survey use ACs for graduate recruitment, external recruitment or internal promotion.

The value of ACs for selection purposes stems in large part from their relatively high criterion-related predictive validity (eg Gaugler et al, 1987). But as ACs came to be more widely used, so was there an increasing recognition of their possible developmental benefits and of certain drawbacks. ACs were seen as having a demotivating effect in some instances, particularly, and not surprisingly, on those who 'failed'. So far as benefits were concerned, there were held to be some for assessors and others for participants; examples of the former are improved interviewing and observational skills which may transfer to normal day-to-day management. For participants, they may obtain a preview of what is involved in/required by the job/role being simulated in the DC. However, Jones and Whitmore (1995: 377) query whether the 'potency' of ACs in changing people's ways of thinking and acting' is actually realised in practice. Moreover, as Thornton (1992)
notes, the demonstrated predictive validity of ACs does not allow us to assume their usefulness for development purposes. Much has been assumed, however, and the perception of the sorts of potential benefits outlined above is one factor which has led to the use of the AC method for developmental purposes.

So far as purposes are concerned, then, practice has broadened. Initially, ACs concentrated on measurement/assessment for selection/promotion. We then saw ‘hybrids’ of various kinds; common here is the AC which is primarily used as a selection device but where feedback might be offered to the participants. This is more likely to happen where the AC is used for internal promotion but is not unknown in the context of external selection. A variant on this theme is the AC that is used for the early identification of potential: although there may be a developmental outcome for individuals, the essential purposes are organisational and it may be that only certain individuals will benefit. It seems likely that there will be some tension between the selection and development functions in all these ‘hybrids’.

There are also DCs which are run on traditional AC lines but, without a selection/promotion decision-making purpose, they are intended to be diagnostic, as summed up in the definition given by Ballantyne and Povah (1995: 150):

The use of assessment centre technology for the identification of individual strengths and weaknesses, in order to diagnose development needs, will facilitate more effective job performance and/or career advancement, which in turn contributes to the attainment of greater organisational success.

Finally, at the other end of the continuum, we have DCs that are designed not just to be diagnostic but also as developmental events in their own right (Thornton, 1992; Jones and Whitmore, 1995), as exemplified by Goodge’s (1991) definition: ‘An offsite process resulting in effective development actions’. Table 1 gives examples of the sorts of outcomes that might be expected of the several variants of DCs.

**TABLE 1 Outcomes of development centres**

<table>
<thead>
<tr>
<th>Development centres aim to assess the skills and competencies of participants in order to do some or all of the following:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Feed back to appesees detailed information on their work behaviour;</td>
</tr>
<tr>
<td>2 Encourage appesees to consider how they can develop competencies they currently lack;</td>
</tr>
<tr>
<td>3 Encourage discussions between appesees and their line managers about the appesees’ future development and the production of a development plan;</td>
</tr>
<tr>
<td>4 Review the overall pool of competencies available to the organisation;</td>
</tr>
<tr>
<td>5 Enable high-level HR or other managers to consider the possible future assignments and training of appesees in the light of the competency profiles derived from development centre activities; and</td>
</tr>
<tr>
<td>6 Identify high potential staff who might be eligible for future fast-track development.</td>
</tr>
</tbody>
</table>

*Arnold, J. 1997. Managing Careers into the 21st Century: 60*

**UNDERLYING ASSUMPTIONS**

Do DCs actually deliver the sorts of outcomes indicated in Table 1? As so little research has been published, we do not have a clear answer; in other words, the developmental and other benefits are assumed and have gone untested. Not only this, a number of other assumptions would appear to underlie the use of DCs. These are often unstated and unquestioned. Drawing on the relevant literature, we identify eight assumptions that fall into three sets:

1 Design, content and operation of the DC
2 Individual motivation, capacity and opportunity to develop
3 Value of planned development and the need for a supportive development environment
1. Design, content and operation of DC

Assumption 1: DCs will give an accurate assessment of performance on particular dimensions pre-determined by job analysis and will provide sufficient information for a clear prescription for development.

As noted earlier, ACs ‘work’, ie their criterion-related predictive validity has been demonstrated. However, we do not know how and why ACs work. One might take the view that, in the context of selection, this does not matter; it is sufficient that ACs predict. Of course this sort of thinking does not hold up to scrutiny from any point of view: eg, from the practitioner point of view we need to understand ‘how’ and ‘why’ in order to develop and improve practice and for DCs this understanding is essential. For selection purposes a pass/fail decision may be enough but if DCs are to serve diagnostic and developmental purposes then participants need specific and detailed feedback. This means that issues to do with construct validity are of heightened importance in DCs (Kauffman et al, 1993). Moreover, issues of content validity cannot be ignored, as the content of the DC – the simulations and tests used – may have a bearing on the acceptance of the process by those going through it.

Construct validity There is extensive literature concerning the problems of establishing the construct validity of ACs. In typical practice an AC (or DC) would measure a number of different ‘dimensions’, that is ‘... a cluster of behaviours that are specific, observable and verifiable, and that can be reliably and logically classified together’ (Thornton and Byham, 1982: 117). These will have been determined through some process of job analysis. What we would expect to find is that assessors’ ratings for any given dimension would be consistent across exercises and would thus be highly correlated; this is convergent validity. Much of the evidence, however, does not support this expectation beyond agreement on some general factor (Schneider and Schmitt, 1992). Indeed, one of the most ‘robust and consistent’ findings about ACs is the existence of the so-called ‘exercise’ or ‘method effect’ (Schneider and Schmitt, 1992). For example, Robertson et al (1987) investigated four separate ACs and found that the inter-exercise correlations for the dimensions measured were low (none greater than .47), averaging (for all dimensions) between .11 and .28. Similarly, McCredie and Shackleton (1994) reported a total mean inter-exercise correlation of .18 for the five dimensions measured in a senior management AC used by Glyndwr International plc.

Not only this, but ACs lack discriminant validity. Thus, it is argued that within an exercise we would expect ratings of the different dimensions not to be correlated. Again, however, expectations are not confirmed. For example, for their four ACs, Robertson et al (1987) found that within-exercise correlations among different dimensions averaged between .49 and .66. In other words, separate, conceptually distinct characteristics of the person are not being measured. Reviewing the evidence on this matter, Kauffman et al (1993) conclude that the construct validity of AC dimensions generally is not supported.

Thornton (1992) has suggested that the relatively high correlation of dimensions within exercises indicates that assessors are unable to distinguish distinct attributes and instead are making situational judgements of performance in each exercise. The evidence of the ‘exercise effect’ leads Robertson et al (1987) to conclude that assessments should be based on exercises designed as work samples. Some empirical support for this view comes from Russell and Dormm’s study (1995), which demonstrated greater construct validity for task-based than traditional trait (dimension) ratings. In other words, ratings tend to be a function more of the exercises used to measure the dimensions than the dimensions themselves. Russell and Dormm suggest that raters are merely judging the quality of performance on each task and not on the subdivided individual constructs/dimensions being rated, which tend to receive similar ratings within an exercise. This suggests that it is difficult to draw any conclusions about whether candidates possess skills which transfer across situations – merely that they were good/bad/indifferent at particular exercises. This implies that the type of exercise has to have a strong similarity to the content of the job for which candidates are being assessed.
Assessments based on tasks or exercises may be appropriate where the purpose is selection/promotion, but are they helpful for development? It seems likely that individuals will want to understand their performance in terms of the skills, abilities etc that they have brought to bear. Thus, if the purpose is to diagnose strengths and developmental needs, then the DC must provide accurate assessments of separate abilities, as feedback based on a deficient diagnosis is clearly likely to be misleading (Thornton, 1992; Kaufman et al, 1993; Joyce et al, 1994). In light of what was said above about the accuracy of assessments made within ACs, we might wonder whether meaningful representations of strengths and weaknesses are generated. That said, summaries (eg Iles and Robertson, 1997) of reaction studies indicate that participants commonly, although not universally, perceive AC assessments to be accurate, furthering their understanding of their strengths and weaknesses.

But what exactly do we mean by assessments in a developmental context? At this point we need to consider the several elements of the ‘assessment’ process within ACs/DCs. It is common (Thornton and Byham, 1982; Blanden and Povah, 1995) to identify a number of steps: observation and recording of behaviour, classification of behaviour, and evaluation/assessment/rating. In some exercises, eg group discussions, the observation and recording will be done by the assessors whereas in others, eg in-trays, observation is not required as the participants themselves record the behavioural evidence. As is shown by the following summary of what is involved in each of these steps, several errors may intrude which reduce the validity of the assessments that are made.

**Observation/recording** Assessors can only observe and record what is on display. We assume first that the simulations have been designed so that they allow the desired dimensions to be displayed. However, as Ahmed et al (1997) have pointed out recently, there is little in the literature to guide the practitioner, yet deficiency in exercise design is likely to introduce a fundamental flaw in the AC/DC. We also assume that participants will take the opportunity to perform. The absence of desired behaviour is difficult to interpret. Is the participant ‘weak’ in the particular dimension? Have they failed to perceive what is required? This last question bears on the issue of ‘transparency’ that is examined by Kleinmann et al (1996: 68) who demonstrated higher construct validity under conditions of transparency:

> Transparency means that the candidates in an assessment centre have been informed about which dimensions and which relevant operationalisations of behaviour are relevant to the assessment in each different exercise.

Kleinmann et al recognise the possibility of ‘play-acting’ but suggest that this may be difficult to sustain across the duration of an AC. Also, where the outcome is intended to be predominantly developmental, it is hard to see that this will be to the individual’s advantage. Furthermore, if transparency does promote construct validity, it seems odd not to disclose the dimensions being assessed, otherwise there is the risk of second-guessing which may lead to the display of the ‘wrong’ behaviour. It might be argued that judging what is required is part of the assessment experience but if this is the case then we would expect this kind of judgement to be one of the dimensions that will be assessed explicitly. Furthermore, exercises need to be designed not merely to allow particular behaviour to be displayed, but rather to require the behaviour to be demonstrated; if it is not, assessors may then draw a negative conclusion rather than no conclusion (Woodruffe, 1990). That said, we need to recognise that for some simulations the opportunity to display relevant behaviour may be limited. For example, a group discussion might last for about 40 minutes. With, say, six participants and assuming equal ‘air time’, this means about seven minutes of behaviour for each person, not all of which will be relevant. Thus, the opportunity to display relevant behaviour may vary from simulation to simulation and, for certain exercises, the evidential base for forming an assessment may be very slender indeed.
We further assume that the assessors will accurately record (ie describe) what they see displayed. As noted above, for some simulations the record is provided by the participants themselves, but for others the recording is done by the assessors and generally it is advocated that the record be made at the time that the behaviour is displayed.

But does the assessor record everything? According to Thornton and Byham (1982: 220), ‘Assessors are encouraged to record as much detail as possible. They record in objective terms the specific things the participant says and does.’ And Woodruffe (1990) suggests that the actions of others, say within a group discussion, should also be noted where they have a significant bearing on a particular participant’s behaviour. Assessors cannot record everything, however: recording is, like observation, a selective process (Boyle et al, 1995). But on what basis is the selection made? Presumably, we would hope that what they have been trained to observe – behaviour relevant to the dimensions being assessed – will have sensitised them. However, there remains the possibility that assessors will continue to hold slightly different interpretations of what is meant by a particular dimension (Kauffman et al, 1993) and that temporary mood states will affect their attentiveness. Also, assessors may vary in observational skill (Zedeck, 1986). We cannot be sure that those who are poorer at this skill will have been selected out by their organisations, hence it is possible that those who are weaker threaten the validity of the assessments made.

Classification of behaviour Thornton and Byham (1982) advocate that the classification of behaviour be done immediately after each simulation. Basically, it involves assigning the observed behaviour to dimensions which, in other words, serve an organising function and in so doing may reduce some of the cognitive demands placed on assessors. In this regard, Zedeck (1986: 28) suggests that dimensions serve as prototypes:

Behaviours observed in assessment centre activities are categorised under particular dimensions because the behaviours contain or reflect most of the expected features to be found in that category. Behaviours are categorised by virtue of the degree to which they are ‘prototypical’ of the dimension in question.

In the case of DCs which may not be simulating a specific target job, the prototype may be of a family of jobs at a particular level, eg middle-management roles.

As Zedeck implies, classification is not a straightforward, mechanical task: it is an inferential and judgemental process (Zedeck, 1986; Ballantyne and Povah, 1995) that is fraught with difficulties. There is the problem, noted above, of the non-display of relevant behaviour. Some behaviour that is displayed may be classifiable in terms of more than one dimension and to this extent the boundaries of dimensions may be fuzzy and overlapping. Other behaviours may not relate to dimensions that are to be assessed. Such behaviour presumably ought to be ignored but there is nonetheless the risk that they will be noted by assessors and so contaminate their judgement.

The behavioural definitions of dimensions are not all-encompassing (Kauffman et al, 1993); that is, we cannot completely specify all the behaviour associated with a particular dimension. One implication of this is that dimensions may hold different meanings for different assessors. Dimensions, especially those of a more global kind, may aggregate different types of behaviour which are similar but not the same. Therefore, behaviour in one exercise on one dimension may be very different from behaviour on that dimension in another exercise. Even if the behaviour does not vary, assessors’ interpretations may do so. All of the above may lead to problems of classification and, in Zedeck’s (1986) view, these are reasons why we should expect dimension intercorrelation within an exercise and low intercorrelations among exercises.

Added to this, there is the view that we might expect a given dimension to be expressed in different ways in different exercises because of the demands of the situations being simulated (Ballantyne and Povah, 1995; Schneider and Schnitt, 1992). In other words, the idea here is that there may genuinely be an exercise effect (Neidig and Neidig, 1984). As Silverman et al (1986: 576) put it:
Some people may simply perform better in some types of exercises, such as
one-on-one role plays, while others may perform better in other exercises, such
as leaderless group discussions.

Assessment There is no set procedure for assessing (ie rating) the categorised behaviour.
Woodruffe (1990), for example, advocates assessment after each exercise on a dimension-by-
dimension basis. Assessors’ individual ratings will then be shared in their final conference where
they try to reach some consensus about each dimension across the different exercises. In the model
advocated by Thornton and Byham (1982) and Thornton (1992), less emphasis is placed on
assessing after each exercise. Rather, there is much fuller sharing of evidence in the final meeting;
for each participant, relevant behaviour is reported on an exercise-by-exercise basis, with ratings
being based largely on what is exposed in this discussion. The little research (eg Silverman et al,
1986; Harris et al, 1993) has shown inconsistent results, so we do not know with any certainty
whether one approach is preferable to the other, but the possibility clearly exists that the procedure
followed may have an impact on the validity of the assessments made (Boyle et al, 1995).

Moreover, in the context of a development centre we might question the value of ratings,
particularly an overall rating. As Fournies (1977) has noted in the context of performance
appraisal, it is far from clear that ratings have much functional value so far as development is
concerned. Therefore, we might take the view that, where there is no promotion/selection
outcome, the final assessment/rating step is not needed. Although participants might be
interested to know how they have been rated, the argument is that this is not especially helpful
information. From the point of view of development, it is the feedback of observed and classified
behaviour which is valuable.

Of course, it is assumed here that such behavioural information is required; this is the position
commonly taken in feedback models eg Latham and Marchbank (1994). Goode (1991) has
suggested, however, that over-complex behavioural descriptions of dimensions need some
unnarralling to produce meaningful development plans.

Content validity Given the ‘problem’ of the exercise effect, several commentators (eg Sackett and
Harris 1988; Crawley et al, 1990; Russell and Domm 1995) have advocated that assessments be
based on exercises rather than dimensions. As indicated above, however, from a developmental
point of view there is the need to provide feedback and it is unlikely to be sufficient for
participants to be told that they have done well on some exercises and less well on others. They are
likely to want to know why, and this implies giving dimensional feedback. But should this be
aggregated across exercises or provided dimension by dimension within single exercises?

If we accept the argument that behaviour is, to some extent, situationally specific, then the latter
would seem to be preferable as such specific, detailed behavioural evidence is likely to be more
developmentally relevant (Boehm, 1985). And, if a dimension is displayed in different ways in
different situations, then we would want to capture as many as possible of those different
situations in the DC and represent them as faithfully as possible. The fidelity of simulations is
important also from the point of view of participant acceptance, ie ‘impact validity’ or ‘social
validity’. However, it is difficult to design exercises which follow a close point-to-point
correspondence with the real workplace when assessing people who are doing similar but not
identical jobs. Close fit of exercises to work will almost inevitably favour one group of people to
the disadvantage of others. One of the strengths claimed for DCs (Blanksby and Iles, 1993) is that
they can bring together people from diverse parts of the organisation and from different
professional disciplines to facilitate networking. This conflicts with the need to produce close-fit
exercises which participants find credible and which may produce sound developmental advice.
Also, Sackett and Ryan (1985) have noted that it is difficult to claim much content validity for a
centre which is looking at potential since, in many cases, people have had little experience in the
areas on which they are being assessed.
Summary We have dealt with the assumption of accuracy at some length because it is the foundation on which the supposed worth of DCs is based. Unfortunately, the assumption is found wanting. Although ACs can possess acceptable levels of predictive validity, their accuracy in other respects has not been conclusively demonstrated. From the point of view of selection decision-making, we might choose to content ourselves with predictive validity. So far as developmental outcomes are concerned, we have the choice of adopting a design for the DC which avoids some of the problems of assessment. But in between the two extremes we have the various hybrid forms and it is with these that all of the problems seem likely to be most acutely felt.

Assumption 2: In hybrid centres designed to be part assessment and part development, the use of results for future decision-making will not interfere with development aims to a significant extent.

There is no doubt that the cost of developing and running ACs/DCs is substantial; this makes it likely that the organisation will want to make the fullest possible use of the information which is generated. This probably is one reason why ‘hybrid’ forms – serving both decision-making and development functions – are so common. Thus, an AC used primarily for selection purposes will often serve development functions of one sort or another. And we perhaps ought not to be surprised if we were to find relatively few examples of DCs yielding only individual development information. Thus, one small study (Jackson and Yeates, 1993) of DCs within 10 UK organisations showed that, while most are intended to assist participants with their development needs, most also have the aim of identifying promotable talent. In ‘hybrids’ it seems likely that organisations will hope that both the selection/promotion and development functions are able to run successfully together.

However, Thornton (1992) has suggested that achieving multiple purposes is very difficult. In many ways the goals are incompatible. Thus, Thornton suggests that different design principles underlie different purposes, with consequent implications for such matters as what is assessed, how the AC/DC operates etc. For example, ACs designed to assess long-term potential may be more concerned to assess individuals’ more stable personal characteristics, whereas in DCs the emphasis may be on skills that are open to modification. Different output is also required; as indicated above, specific and detailed information is needed by the individual for developmental purposes and organisational needs may be satisfied with lesser information. If the procedure is run on traditional AC lines, it is possible that fuller details will not be generated and there may also be the problem of the exercise effect.

It also has been suggested (Smith and Robertson, 1992) that DCs require an environment where it is acceptable to make mistakes. This will always be a problem in hybrid centres. There is, too, the matter of how the information is used; if it is for decision making then it can be difficult to draw out a meaningful discussion on development. And if those who do ‘better’ – e.g. are seen as having greater potential receive ‘better’ treatment in terms of development opportunities offered – then there is the risk that the procedure will simply be seen as a selection centre. As well as Thornton, others (e.g. Griffiths and Goadge, 1994) have suggested a different design if the main purpose is to be developmental. This should include on-going feedback, peer review etc along the lines of a career counselling session. Williams and Dobson (1993) and Iles and Forster (1994) have also advocated more collaborative designs and the examples that the latter provide show that such DCs can be organisationally feasible. But, equally, there will be many organisations in which such designs, or different designs for different purposes, simply will be seen as unrealistic because of the costs involved. ‘Hybrids’ are likely to continue, therefore, and many of the sorts of tensions outlined above will remain unresolved.

2. Individual motivation, capacity and opportunity to develop

Assumption 3: Development centres will increase motivation to develop skills/career. That individuals participate in DCs might lead us to assume that they want to develop – in other words, they already are motivated. This raises the issues of how and why participants come to be
nominated/selected. In the 10 organisations studied by Jackson and Yeates (1993), line manager nomination was common, provision for self-nomination less so. The larger Industrial Society (1996) survey also indicates that participant selection by the organisation is more common than self-nomination.

Although this much is known about nomination, we know much less about the underlying motivations. Is need for achievement the driving force? A perceived need to improve one’s skill or ability? A more general belief in the value of development? We know little about the antecedents which lead people to seek out development opportunities (Tharenou, 1997a, b). Moreover, given the prevalence of organisational nomination, we should not assume that all those who attend are wholly willing participants. Although many are likely to have been nominated because they are already seen as having potential for advancement, for others the nomination may have some ‘remedial’ intent. At the very least it seems reasonable to assume that participants’ levels of motivation will vary – not only their motivation to attend but also their more general motivation to develop themselves. And as Boehm (1985) has pointed out, it is merely an assumption that people will be sufficiently motivated to undertake appropriate development activities. It seems fairly well-established now that motivation to learn and develop and actual learning are fairly strongly related (e.g. Noe and Schmitt, 1986; Hicks and Klimoski, 1987; Mathieu et al., 1990). Also, some research suggests that personal factors such as career motivation are associated with taking developmental action and advancement (Jones and Whitmore, 1995). This suggests that those are who already motivated to develop might derive more benefit than those lacking such motivation. However, there is the possibility that involvement in the process of identifying development needs may enhance motivation to learn (Baldwin et al., 1991; Noe and Wilk, 1993). The question then is whether participants will be more willing to engage in developmental activities after the DC. Baldwin et al.’s research suggests participant involvement may have the hypothesised motivating effect provided that action is taken to meet the needs identified; in their study this was the delivery of appropriate training. Participant involvement may also have other benefits. For example, Noe and Schmitt (1986) studied participant responses to a training programme in which needs were first identified via an AC. Trainees who reacted positively to the needs-assessment procedure were more likely to be satisfied with training programme content:

It appears that trainees who agreed with the assessments of the skill weaknesses on which their training assignment was based were more likely to perceive the content of the training programme to be useful and helpful for skill improvement.

Noe and Schmitt, 1986: 517

Noe and Schmitt also found some association between career planning activity and behaviour change (both self-reported). But, quite apart from whether motivation is enhanced, it is probable that expectations will be raised and this was a concern identified in the Industrial Society (1996) survey. So far as the individual is concerned, frustration may result if needs are not met, perhaps leading to demotivation. From the organisational perspective, raised expectations have to be met and there is a cost attached to this. But there would seem to be little point in embarking on an intervention such as a DC if there is not a willingness to provide the later support that will be required.

So far as DCs are concerned, we have little relevant evidence about their motivational impact. What we do know tends to be about ACs and is only tangentially related to motivation to develop. For example, we know that, generally speaking, the reaction is positive, although Dodd (1977) found that between 10 and 30 per cent of participants believed that ACs did not accurately assess strengths and weaknesses. This is likely to inhibit motivation to develop. However, perceptions also depend on the outcome for the individual – did they pass or fail? Thus, Teel and Dubois (1983) discovered that high scorers tended to have more positive views of the AC than did low scorers who perceived it to be less fair and accurate and who also reported fewer career benefits. And in a study of two ACs, Robertson et al. (1991) found that those who passed perceived greater positive career impact as compared to those who failed.
Clearly, then, the impact for some may be negative; Jones and Whitmore (1995) discovered that more developmental recommendations were made for those who received lower ratings, and the more recommendations, the less there was acceptance of feedback, despite assurances that recommendations were purely developmental. Noe and Steffy (1987) also found that those who receive the most negative evaluations may be less likely to experiment with managerial skills; they suggest the reason for this may be that the centre serves as a job preview and lowers expectations among those with low evaluations, resulting in lower exploration in their current position. Thus, the demotivational effect of performing badly within the centre may mean that an individual’s knowledge of their strengths and weaknesses may lead to less developmental activity. In some cases participants may use the results to redefine their career and leave the organisation (Iles et al., 1989). This could be a reaction to a perception of poor fit between them and the employer and the belief that any development plan is unlikely to resolve that lack of fit. Choosing to leave one’s organisation is, of course, development activity for the individual. Whether it is what the employer wants is another matter – much depends on who is leaving!

Although the diagnosis that takes place within the DC may be the starting point for developmental activity, Dubin (1990) suggests that the decision to participate in that activity is likely to be influenced by beliefs that it will result in favourable outcomes. In other words, the DC may indicate a need for skill enhancement; the extent to which this is acted on may depend on the potential benefits – or non-benefits of not taking action – and their attractiveness for the individual (Maurer and Tarulli, 1994). The provision of realistic information about the features and benefits of development may increase motivation to learn (Hicks and Klimoski, 1987). This points to the need to consider DCs within a broader framework of career development policy; the climate of the organisation must support development and the opportunities must exist. It is important to recognise also that the decision to take part in development activity depends on a range of other factors – some organisational, some individual (Tharenou, 1997a, b) – beyond the results of the AC/DC.

In summary, it seems that there is a paradox here. Those who do well on DCs may be seen as having fewer development needs – as compared to those who do less well – such that development action is not seen as necessary or beneficial. However, where the benefits are attractive and valued, action may take place: in this regard, clear and specific development goals may serve a motivating function. In particular, those seen as having greater potential for advancement stand to gain more, may be motivated to act and may indeed be offered more development opportunities. In contrast, those who perform less well may have greater development needs – in some cases so many that they may be less motivated to develop their skills. All this suggests that for some participants there is a risk of DCs having the opposite effects from those intended.

Assumption 4: Identified areas for improvement are perceived to be important for future success.

Important in whose eyes? There are two main perspectives: that of the organisation and that of the individual. So far as the organisation is concerned, we might assume that the dimensions being assessed are important for the future success of the organisation. This implies that the job/role analysis undertaken will have been forward-looking in nature (Schneider et al., 1997; Visser et al., 1997). The organisation must also see investing in development as worthwhile and as something that is associated with its future success; only then is it likely that recommendations from the DC will be given sufficient support for the development plans to have much effect.

Individuals, however, are likely to be more interested in their own future success. So, if areas for improvement are identified, individuals need to see that there is something in it for them to participate in developmental activity. In other words, as noted earlier, such participation is likely to be influenced by beliefs that the activity will result in personally favourable outcomes (Dubin, 1990), whether these be of an extrinsic (promotion) or intrinsic (interest/challenge within the job)
kind. We would expect that individuals must perceive a direct link between the skills to be developed and some future outcome that they value. Expectancy theory (Vroom, 1964) and much of the literature on training (eg Gilbert, 1982; Noe, 1986) indicate that it is likely that the outcomes of training must be judged as relevant or instrumental to job performance and that better job performance must have some value for trainees. In a study looking specifically at a DC for managers in a large insurance company, Jones and Whitmore (1995) found that developmental recommendations tended to be followed only if rating feedback sent a positive message about future advancement. This reinforces the importance of placing DCs within a broader developmental climate and not regarding them as one-off events.

**Assumption 5: Individuals are able to develop skills in areas of weakness.**

In broad terms, there would seem to be three main aspects to this. First, as already discussed, individuals must have the requisite motivation. Secondly, they must have the opportunity to develop they have to have appropriate development experiences. Thirdly, individuals must have the ability; this bears on the issue of what is assessed and the assumption that a DC will assess properties of the person that are capable of being developed. Is this what DCs do?

Some research (eg Bray and Howard, 1983; Jones and Whitmore, 1995) has demonstrated that it is the assessment of relatively stable attributes that is important for the predictive accuracy of ACs. A notable study (Englebrecht and Fischer, 1995) reports a developmental AC (plus subsequent feedback and action planning) as having a positive impact on work performance. However, what was especially notable was that the least impact was on those attributes less susceptible to change – in this instance, those of a cognitive nature.

This raises the issue of what we understand by development. If our interest is in changing skills, then these are what we should seek to assess in the DC. In doing so, however, we need to recognise the possibility that the more malleable skills, although possibly important for current performance, may not be the ones associated with future potential and advancement (Jones and Whitmore, 1995). If it is the latter that we understand by development, then our interest should be in assessing more stable attributes of the person. These may be less susceptible to change other than perhaps through the most concerted of efforts (Thornton and Byham, 1982; Woodruffe, 1990).

In other words, different purposes may require us to assess very different attributes of the person. This brings us back to the problem identified earlier, that of accommodating different purposes within a single design.

3. Planned development and the supportive environment

**Assumption 6: An active or planned approach to self-development will enhance an individual’s skills/promotability/career prospects.**

The term ‘self-development’ may be interpreted in various ways. Burgoyne et al (1978) list 10 possible interpretations, although these fall into a smaller number of categories. Thus, there is the distinction, albeit not hard-and-fast, between development of self – associated with ideas such as self-actualisation and psychological growth – and development by self (Boydell, 1985). The latter is often associated with acceptance of personal responsibility for learning and career development; that is, the idea of self-directed or self-managed learning. In practice, acceptance of such responsibility involves various activities, such as obtaining information about strengths and weaknesses, career information, developing goals and plans to achieve them, actually implementing the plans (taking developmental action), and so forth. All these things would be seen as self-initiated. Moreover, it is expected that such activity is ‘better’ than passivity; those with clear and specific career goals are more likely to achieve them than those without, as we might expect from goal-setting theory.

Taking more of an organisational perspective, Arnold (1997) has noted the term self-development has taken on two main meanings. On the one hand, there is an aggressive ‘You’re on
your own' view. The second view, while placing responsibility on the individual, sees organisations as recognising that they share in the responsibility, at least to the extent of providing supports of one kind or another (Hirsh and Jackson, 1996). These supports include a wide range of interventions – Arnold lists 14 – which may serve many functions.

One such intervention is the DC which, as Table 1 indicated, may have several purposes, including the development of a plan of action to try to improve areas of weakness and build on strengths that have been diagnosed. In other words, the diagnosis from the DC leads to a plan of action, this action in turn leading to change for the better for the individual. What evidence is there that this planned approach will lead to increased skill or career development? There is remarkably little evidence about the effect of career development interventions (Russell, 1991; Arnold, 1997); rather, we have a lengthy list of largely unsupported claims.

One study (Noe, 1995: 129) indicated some association between exploratory behaviour and preparedness for development:

The more employees sought information from peers, friends and managers regarding skill strengths and weaknesses and potential career areas, the more willing they were to participate in development activities.

However, whether or not that willingness converts to action depends on a whole range of other factors, some organisational and some individual. Noe and Wilk (1993), for example, have pointed to the possible moderating effect that self-efficacy may have on development activity; those high in self-efficacy are more likely to take responsibility for personal development. However, action will also depend partly on what the organisation allows or encourages (Maurer and Tarulli, 1994). This takes us on to a further assumption.

**Assumption 7: Active or planned development is compatible with organisational constraints and culture.**

When we refer to a DC, we probably mean a collection of related activities: assessment, diagnosis of development needs, feedback and a development plan. Moreover, if DCs are to have a longer-term effect, the plan has to be implemented and a follow-up mechanism put in place (Jackson and Yeates, 1993). In other words, a DC is not an isolated event and can only function if it is part of a wider process which includes support for individual development back in the workplace (Seegers, 1989). Generally speaking, it is assumed that such a supportive environment exists. Yet, as Arnold (1997) has indicated, the self-development philosophy can in practice mean the abrogation of all responsibility on the organisation's part. We would expect this to be a barrier to development. As Tharenou (1997a: 21) puts it:

Work environment factors that should be most related to participation in development would be those that make it possible to attend and that reduce barriers and situational constraints.

In an evaluation of a career development programme based around a personal development plan (PDP), Giles and West (1995) discovered that many attempts to implement the plan came across difficulties with the economic state of the organisation and the indifference toward career plans by supervisors. It seems unlikely that any development will take place if all parties do not see it as a valuable process on which time should be spent. Various commentators (eg Noe and Ford, 1992) have attached particular importance to the role of the line manager. Thus, Noe (1996: 123) suggests that employees are likely to be:

More motivated to engage in career management activities when their manager encourages discussions related to development and career issues, is willing to identify resources to help the employee with specific problems, and assists the employee in setting performance and career goals.

Some of his research (Noe and Wilk, 1993) indicates that the perception of line manager support does influence development activity. Also, Francis-Smythe and Smith (1997) report an association
between perceptions of line manager support and performance on an AC. Those who did well were more likely to be those who saw themselves as receiving good support. Research by Maurer and Tarulli (1994) further provides some evidence of an association between supervisor support and participation in development activities. However, they noted that such support needs to be valued by employees.

Yet supervisors may be unwilling to assume these sorts of responsibilities (Gutteridge, 1986) and may lack the requisite skills and knowledge. In this sense, managers themselves may feel unsupported (Hirsh and Jackson, 1996). But, given the apparent importance of supervisor support, Noe (1996: 123) suggests that:

To facilitate development behaviour organisations might best use their resources to train managers in skills needed to support employee development... rather than purchasing or developing career management programmes that focus on career exploration and career strategies.

It therefore seems likely that a DC will only have a positive effect if it is an integral part of a concerted effort aimed at fostering development within the organisation. In short, organisational policies and practices have to provide support for employee development (Maurer and Tarulli, 1994; Tharenou, 1997a, b).

**Assumption 8:** We know which development activities will have an impact on particular skills.

We would expect that development plans will include some statement of development needs, eg in terms of skills, along with some specification of the activities to be undertaken that will bring about the desired development. But how well do we know what sort of development activity will best meet a particular need? We know that much development takes place within the context of doing a job (Nicholson and West, 1988), although little is known about the developmental components of jobs (Noe and Ford, 1992; McCauley et al, 1994). Also, there seems to be considerable doubt as to the role that individual differences play in the determination of suitable development activities. Much of this has centred on individual learning styles (Honey and Mumford, 1986) but there seems to be little known about the effectiveness of different learning styles (Keefe, 1987).

Generally speaking, we have a poor understanding of what sorts of needs on the part of what sorts of individuals are best met by a particular kind of development activity. Assuming there is a feedback discussion at which development needs and actions are considered, then it would seem that the knowledge and skill of the feedback giver will be of particular importance. Moreover, we cannot ignore that the organisational context opportunities to develop must exist and resources need to be available.

**CONCLUSIONS**

In this article we have sought to ask many questions about DCs. Do they diagnose development needs? Do they lead to development action? Do they lead to changes in behaviour and performance? Unfortunately, clear-cut answers to such questions are not available. Rather, the supposed benefits and the quite extensive use of DCs depend on a series of untested assumptions. Such evidence as exists is not especially encouraging. The popular application of the AC method for development purposes stems in large part from the demonstrated predictive validity of ACs. Although the value of the method for selection/promotion purposes is not in question here, there are doubts about aspects of AC validity, especially construct validity. Do ACs yield valid assessments of the individual dimensions of performance that they set out to measure? This is in doubt. But for development purposes what we need is valid information about participants' capabilities.

As we noted earlier, where the outcome is simply a selection decision, we might choose to not be especially concerned about the apparent lack of construct validity. We might content ourselves
with the knowledge that ACs predict future potential and performance. However, this rather ignores business pressures to try to wring as much as possible from a process that is costly to develop and operate. There is also the desire on the part of those who are assessed to receive feedback on how they have done. Hence, we commonly find that ACs will often have a secondary development purpose, even if the primary outcome is selection/promotion.

By the same token, it is not uncommon for DCs to have a secondary purpose, although this may be rather more subtle. For example, a DC may have the aim of diagnosing training and development needs and identifying future potential. Such a DC could very easily have a secondary selection purpose, as those who are seen as having potential may be the ones who get the development opportunities. Indeed, some of the literature that we have reviewed suggests that, for some participants, the effect of DC participation may be the exact opposite of what is intended. Thus we need to recognise the possibility of DCs having a demotivating effect, particularly on those seen as having many development needs. We know little about why people take part in DCs, yet this is important as participant motivation may be one of the factors that will influence what they get out of the experience.

Ideally, perhaps, we might wish to see very different designs for DCs run along collaborative lines. Such designs are feasible and do exist, but many organisations will see them as expensive luxuries. In our view, therefore, 'hybrid' forms aimed at serving dual selection and development purposes are likely to remain commonplace. But such hybrids are riddled with tensions. For example, the different purposes may require the assessment of different kinds of attributes - more stable ones, such as underlying cognitive capacities, in the case of selection/promotion, but more malleable ones where skill/behaviour change is the intended outcome. In a DC we might not need assessment (in the sense of rating) at all - specific, detailed feedback about behaviour may be the primary requirement. In short, the expectations that we have of hybrid DCs (or ACs) may be far too high: the very real risk is that neither purpose will be served well.

Finally, it is worth noting that other interventions might be used alongside or instead of DCs to diagnose development needs. For example, performance management systems (Williams, 1998) might serve both to diagnose needs and facilitate development through developmental objectives. Multisource feedback has become popular today (Williams, 1998) and much has been claimed for this tool, but here again the hard evidence is lacking (Fletcher and Baldry, 1999). And there is a range of other tools such as workbooks, workshops, questionnaires etc (Arnold, 1997; Kidd, 1997). But it seems likely that much of what we have said about DCs applies to these also. In particular, it seems unhelpful to regard diagnostic interventions as stand-alone activities. The supportiveness of the organisational environment, including supervisor support, is of special significance since, if this is lacking, there is every likelihood that any positive effect that a DC (or workshop, or multisource feedback) might have will quickly dissipate. If there is to be any prospect of DCs delivering benefits, organisations have to be willing to follow up and deliver on the expectations that have been raised, and this requires both policies and skills for promoting employee development.

REFERENCES

Boehm, V. 1985. 'Using assessment centres for management development: five applications'.

HUMAN RESOURCE MANAGEMENT JOURNAL - VOL 9 NO 2
Development centres – a review of assumptions

Wiley.
Assessment Center Operations, Alexandria, VA:IPMA.
261. Falmer, Sussex: Institute of Manpower Studies.
Jones, R. G. and Whitmore, M. D. 1995. ‘Evaluating developmental assessment centers as
1, no. 4, 213-223.
School Principals.
Kidd, J. M. 1997. ‘Assessment for self-managed career development’ in International Handbook of
Selection and Assessment. N. Anderson and P. Herriot (eds). Chichester: Wiley.
Kleinnmann, M., Kuptsch, C. and Killer, O. 1996. ‘Transparency: a necessary requirement for the
1, 67-84.
McCaulley, C. D, Ruderman, M. N, Ohlott, P. J. and Morrow, J. E. 1994. ‘Assessing the develop-
McCredie, H. and Shackleton, V. 1994. ‘The development and interim validation of a dimensions-
91-101.
Mathieu, J. E, Tannenbaum, S. I. and Salas, E. 1990. ‘An influence of individual and situational
Maurer, T. J. and Tarulli, B. A. 1994. ‘Investigation of perceived environment, perceived outcome
and person variables in relationship to voluntary development activity by employees’. Journal of
Cambridge: Cambridge University Press.
Noe, R. A. and Steffy, B. 1987. ‘The influence of individual characteristics and assessment center
evaluation on career exploration behaviour and job involvement’. Journal of Vocational Behaviour,
Robertson, I. T., Gratton, L. and Sharpley, D. 1987. ‘The psychometric properties and design of
managerial assessment centres: dimensions into exercises won't go'.  

The impact of personnel selection and assessment methods on candidates'.  
Human Relations, Vol. 44, no. 9, 963-981.

Russell, J. E. A. 1991. 'Career development interventions in organisations'.  

Russell C. J. and Domm, D. R. 1995. 'Two field tests of an explanation of assessment centre validity'.  

Sackett, P. and Harris, M. 1988. 'A further examination of the construct underlying assessment center ratings'.  

Sackett, P. and Ryan, A. 1985. 'A review of recent assessment centre research'.  

Schneider, B., Kristof-Brown, A. L., Goldstein, H. W. and Smith, D. B. 1997. 'What is this thing called fit?' in  
International Handbook of Selection and Assessment. N. Anderson and P. Herriot (eds).  
Chichester: Wiley.

Schneider, J. R. and Schmitt, N. 1992. 'An exercise design approach to understanding assessment center dimension and exercise constructs'.  


Tharenou, P. 1997b. 'Managerial career advancement'. International Review of Industrial and Organizational Psychology, Vol. 12, 39-93.


Williams, A. P. O. and Dobson, P. 1993. 'The use of developmental assessment centres on MBA programmes'.  

Williams, R. S. 1998. Performance Management: Perspectives on Employee Development, London:  
International Thomson Business Press.
