Journal of Industrial Psychology, 1989, 15(2), 1-10
Tydskrif vir Bedryfsielkunde, 1989, 15(2), 1-10

VALIDATION OF AN ASSESSMENT CENTRE AGAINST BARS: AN EXPERIENCE WITH PERFORMANCE RELATED CRITERIA

H.H. SPANGENBERG

J.J. ESTERHUYSE

J.H. VISSER

J.E. BRIEDENHANN

Stellenbosch Farmers' Winery
Stellenbosch

C.J. CALITZ

Department of Industrial Psychology
University of Stellenbosch

ABSTRACT

The objective of this study is to determine the validity of a middle management assessment centre. Behaviourally Anchored Rating Scales were developed specifically for this purpose. The results testify to the validity of this assessment centre and lend support to the argument that assessment centres do indeed predict performance criteria (viz a viz potential/advancement criteria). The validity coefficients for the various dimensions are indicated.

OPSOMMING

Die doel van hierdie studie is om die geldigheid van 'n takseersentrum vir middelbestuur te bepaal. Vir dié doel is gedragsgeankerde skale as kriterium spesiaal ontwikkel. Die resultate dui op die geldigheid van hierdie takseersentrum en verleen addisionele steun aan die argument dat takseersentrums wel prestasiekriteria (viz a viz potensiaal/bevorderingskriteria) voorspel. Die geldigheidskoëffisiënte van die onderskeie bestuursdimensies word aangetoon.

Assessment centres were introduced at Stellenbosch Farmers' Winery in September 1979 with the objective of identifying overall management potential and assessing management skills of individual managers. They also formed an integral part of the company's manpower planning and development effort. The approach has been to start at the top and work downwards, which proved to be the most appropriate one for the company, since it contributed to the acceptability and credibility of the assessment centre concept. This credibility ensured the increased use of centre results as an additional input for promotions and transfers.

When most senior managers had attended the Senior Assessment Centre, a needs analysis exercise was conducted at senior management level to determine the nature and exact level of the proposed assessment centre for middle management. Regional groups of senior managers unanimously recommended that a general management assessment centre be constructed for levels just below the running centre.

The objectives of the proposed Middle Management Assessment Centre (MMAC) were to provide a comprehensive assessment of the managerial skills of senior middle managers as well as to determine the managerial readiness of more junior managers and professional people aspiring to middle management positions. Assessment data was to be used both as an input for compiling personal development plans and for the purpose of developing management succession charts.

Since assessment data would be used for making promotional decisions, the assessment instrument had to be a valid one. From the outset, therefore, the intent was to construct an assessment centre that would be able to predict relevant performance criteria.

Requests for reprints should be addressed to H.H. Spangenberg, Stellenbosch Farmers' Winery, P.O. Box 46, Stellenbosch, 7600

CONSTRUCTION OF THE MMAC

The construction model for the MMAC was based to a large degree on the work of Len Slivinski of the Canadian Public Service (Slivinski, 1977; 1981), assisted by Lois Crooks of Educational Testing Services, Princeton. It was decided to use this model because of the significant improvements Slivinski and his colleagues made to the traditional assessment centre technology. These include the principles of central character and continuity of content material; the use of professional role players in interactive exercises; and more emphasis on assigning a leadership role in a group discussion.

The steps involved in the construction and validation of the MMAC are shown in Figure 1. As can be seen in steps 1 to 3 the MMAC is based on comprehensive analysis of the target level, which includes the content as well as the context of jobs at that level. The column "People Involved" indicates the extent to which directors, senior managers and a Management Review Committee, who monitored the construction process by supplying so-called expert judgement (Slivinski, 1981, p. 18), were involved. For a detailed discussion of the construction and implementation of the MMAC see Spangenberg, (in press).

Following the job and organizational analyses, 13 dimensions were formulated and five exercises constructed. The Verbal Fact Finding exercise was adapted from an exercise developed by Development Dimensions International (DDI).

An indication of the extent to which dimensions are covered by the various exercises can be seen in Table 1.

It was decided to use Anchored Rating Scales (BARS) as criteria for the following reasons:

Firstly, formal and informal surveys of results of the current SFW Performance Appraisal System have indicated that in some cases line managers were abusing the system by rating

FIGURE 1
STEPS IN CONSTRUCTING AN ASSESSMENT CENTRE.

STEPS	INSTRUMENT/METHOD	PEOPLE INVOLVED
1. NEEDS ANALYSIS	1. GROUP SESSIONS	1. DIRECTORS & SENIOR MANAGERS WHO ATTENDED SENIOR ASSESSMENT CENTRE
2. JOB ANALYSIS	1. REVIEW OF LITERATURE	1. TARGET GROUP MANAGERS
	2. "JOB ANALYSIS INTERVIEW GUIDE"	2. TARGET GROUP MANAGERS
	3. "DAILY ACTIVITY LOG" (DIARY LOG)	
3. ORGANIZATIONAL CLIMATE ANALYSIS	1. "ORGANIZATIONAL ANALYSIS QUESTIONNAIRE" (BECKHARD)	1. TARGET GROUP MANAGERS
	2. ORGANIZATIONAL CLIMATE QUESTIONNAIRES	2. TARGET GROUP MANAGERS
	2.1 Litwin & Stringer 2.2 Schneider & Bartlett	
4. IDENTIFY MANAGERIAL DIMENSIONS		
4.1 Construct & administer schedule of major tasks, problems & managerial skills	1. "Schedule of major tasks, problems & managerial skills"	1. Target group managers – monitored by their immediate superiors
4.2 Identify major task clusters	1. Completed schedule of tasks, problems & managerial skills	
4.3 Construct & administer Problems Questionnaire; process questionnaire; identify most critical problems	1. Questionnaire entitled "Problems Handled by Middle Management"	1. Directors & senior management (Gr 14 - 16) for completion of questionnaire
	2.1 "Problems in priority order"	
	2.2 "Priority Problems as per Major Task Cluster" (Group Discussion)	2. Management Review Committee to help identify most crucial problems
4.4 Collate information & work through all documents with Management Review Committee	1. Summary of sub-parts of "Job Analysis Interview Guide"	1. Management Review Committee
	2. Summaries of Daily Activity Logs	
	3. Summary of Organizational Analysis Questionnaire (Beckhard)	
4.5 Finalize dimensions	1. Literature review	
	1.1 Compile list of Assessment Centre Dimensions	
	2. Plot frequency of dimensions	
	3. Collate information	
	4. Formulate tentative dimensions	1. Management Review Committee
	5. Finalize dimensions	2. Senior Assessment Centre Consultant
5. DECIDE ON ASSESSMENT EXERCISES		1. MANAGEMENT REVIEW COMMITTEE
		2. SENIOR ASSESSMENT CENTRE CONSULTANT
6. CONSTRUCT EXERCISES		1. MANAGEMENT REVIEW COMMITTEE
		2. SENIOR ASSESSMENT CENTRE CONSULTANT
7. IMPLEMENT ASSESSMENT CENTRE		1. SENIOR MANAGERS
8. VALIDATE ASSESSMENT CENTRE	1. BEHAVIOURALLY ANCHORED RATING SCALES (BARS)	1. TARGET LEVEL MANAGERS

FIGURE 1. Steps in Constructing an Assessment Centre.

subordinates abnormally high. This led to some doubt about the validity of the appraisals and it was decided, therefore, not to use the performance appraisals for validation purposes.

Secondly, the need was felt for a behaviourally orientated rating scale that could be used to measure *on the job performance* of assessment centre participants. It is important to note that assessment centres measure only those dimensions which can be effectively *observed* in a simulated situation. In order to measure the impact of assessment centre participation and follow-up development activities on participants it was essential to measure *total work performance.* The BARS used to measure this total performance therefore include variables which could not be measured in the assessment centre.

BARS were constructed in a way similar to the traditional method (Smith & Kendall, 1963). A comprehensive description of the construction of BARS used in this study is published elsewhere (Spangenberg, Esterhuyse, Visser, Briedenhann & Calitz, 1989).

What is relevant here, however, is the fact that BARS were constructed *totally independently* of the job analysis conducted prior to the construction of assessment centre exercises. As a matter of fact they were developed only a year after implementation of the assessment centre. Groups of senior managers who knew the middle management positions intimately were involved on two occasions: firstly, in brainstorming the knowledge, skills, abilities and personal attributes

TABLE 1
MIDDLE MANAGEMENT ASSESSMENT CENTRE: EXERCISES AND DIMENSIONS

EXERCISES[$]					DIMENSIONS	MANAGEMENT AREAS
BI	IB	VFf	GD	IS		
			X	X	Flexibility	PERSONAL
X					Personal Development	DIMENSIONS
	X	X			Analytical Skills	DECISION-MAKING
	X	X			Judgement	
	X			X	Decisiveness	
	X				Planning & Organizing	ADMINISTRATIVE
	X				Delegation	SKILLS
	X				Control	
	X		X	X	Interpersonal Sensitivity	LEADERSHIP
			X		Group Leadership	
				X	Individual Leadership	
				X	Oral Communication	COMMUNICATION
		X			Presentation Skills	

[$] BI = Background Information
IB = In-Basket
VFf = Verbal Fact-finding
GD = Group Discussion
IS = Interview Simulation

(KSAP's) required for successful performance at this level together with behaviour examples of good, average and poor performance on these qualities; and secondly, in rating of behaviour examples during later stages of the construction of BARS. Furthermore, in no way was it attempted to develop rating scales directly comparable to the MMAC dimensions.

The salient characteristics of BARS, i.e. their measuring of total performance, their construction totally independently of assessment centre exercises, and the lack of comparability between rating scales and assessment centre dimensions could possibly result in lower validities but on the other hand increase their ultimate usefulness as measuring instruments.

METHOD

It was decided to use the correlational method since the relationships between a large number of predictor and criterion variables were investigated.

Hypothesis

The main hypothesis for this study is that significant correlations exist between behaviour measurements resulting from the assessment centre and Behaviourally Anchored Rating Scales (BARS).

Sample

The sample consisted of 110 middle and first line managers (equal to Peromnes grades 6 to 8), who attended the MMAC over a period of two years.

Variables

The predictor variables consisted of five biographical items – i.e. Grade, Age, Language, Sex and Race – and thirteen assessment centre dimensions.

TABLE 2
BEHAVIOURALLY ANCHORED RATING SCALES

SCALES	CATEGORY
Self-confidence, Tenacity, Drive, Initiative, Personal involvement, Adaptability/Flexibility, Loyalty/Integrity/Personal Values	PERSONAL QUALITIES
Analytical Skills, Judgement, Organizational & Environmental Awareness & Sensitivity, Decisiveness	DECISION-MAKING SKILLS
Leadership, Interpersonal Relations, Negotiating Skills, Development of Subordinates	LEADERSHIP SKILLS
Planning & Organizing, Delegation, Control	ADMINISTRATIVE SKILLS
Communication	COMMUNICATION SKILLS

The effectiveness criteria included 19 BARS ratings and a total rating (hereafter referred to as Rating) derived from simple addition of the criterion scores.

The BARS can be classified into five categories (see Table 2), which roughly correspond to the dimensions of the MMAC.

Procedure

Using the SPSS computer programme, means, standard deviations and intercorrelations between all variables were calculated.

Both stepwise and simple regression analyses were conducted for individual predictor variables and Rating and separately for groupings of predictor variables and Rating. Pearson correlations between groupings of predictor variables, corresponding groupings of criterion variables and Rating were calculated.

Finally, canonical correlations between various linear combinations of predictor and criterion variables were computed.

RESULTS

Distribution of Predictor and Criterion Variable Scores

Means and standard deviations for all variables are presented in Table 3.

TABLE 3
MEANS AND STANDARD DEVIATIONS OF PREDICTOR AND CRITERION VARIABLES (N = 110)

PREDICTOR VARIABLES			CRITERION VARIABLES		
Variable	**Mean**	**S.D.**	**Variable**	**Mean**	**S.D.**
			Self-confidence	3,67	0,69
Biographical[a]			Tenacity	3,71	0,65
			Drive	3,71	0,71
Age	37,27	7,45	Initiative	3,62	0,71
			Personal Involvement	3,69	0,59
Assessment Centre			Adaptability/Flexibility	3,77	0,55
			Loyalty/Integrity	4,02	0,63
Flexibility	101,25	32,42			
Personal Development	66,36	45,91			
Analytical Skills	89,38	43,37	Analytical Skills	3,55	0,74
Judgement	118,60	26,31	Judgement	3,63	0,58
Decisiveness	126,15	28,20	Organizational & Environmental Awareness & Sensitivity	3,63	0,61
			Decisiveness	3,51	0,61
Planning & Organizing	81,70	46,83	Planning & Organizing	3,57	0,74
Delegation	56,89	39,85	Delegation	3,42	0,67
Control	41,74	45,04	Control	3,55	0,70
Interpersonal Sensitivity	108,80	36,02	Leadership	3,53	0,59
Group Leadership	86,97	46,60	Interpersonal Relations	3,69	0,69
Individual Leadership	82,11	46,45	Negotiating Skills	3,59	0,72
			Development of Subordinates	3,30	0,63
Verbal Communication	112,45	35,34	Verbal Communication	3,70	0,65
Presentation Skills	70,76	36,92			
			Rating	3,62	0,45

Note: The converted mean score for an Assessment Centre dimension is 100.00 and for a criterion rating scale is 3.00

[a] No means or standard deviations were calculated for grade, language, sex or race.

An interesting feature of the above results is that of the thirteen assessment centre dimensions, three which are scored slightly differently from the rest have the lowest standard deviations, i.e. Flexibility (SD = 32,42), Judgement (SD = 26,31), and Decisiveness (SD = 28,20). In scoring these three dimensions a "3" (equivalent to 100) rating is awarded unless either sufficient positive or negative evidence is produced to take the score away from a "3".

A very low mean score for Control was expected since setting controls in the In-Basket seems to have been a problem to many participants.

Intercorrelations between Predictor and Criterion Variables

Pearson correlations between all variables were calculated and significant correlations are tabled in Appendix 1.

An overview of this matrix shows high intercorrelations between criterion variables. This indicates a degree of overlap between the rating scales and that, to some degree, these scales explain the same variance. Correlations are, however, not high enough to detract from the independence of the scales.

Significant correlations between the predictor and criterion variables are shown in Table 4.

TABLE 4
CORRELATIONS BETWEEN PREDICTORS (BIOGRAPHICAL DATA & ASSESSMENT CENTRE DIMENSIONS) AND CRITERIA (BARS & RATING) (N = 110)

PREDICTOR VARIABLES \ CRITERION VARIABLES	ANALYTICAL SKILLS	JUDGEMENT	ORGANIZATIONAL & ENVIRONMENTAL SENSITIVITY	DECISIVENESS	PLANNING & ORGANIZING	DELEGATION	CONTROLLING	LEADERSHIP	INTERPERSONAL RELATIONS	NEGOTIATING SKILLS	DEVELOPMENT OF SUBORDINATES	COMMUNICATION	SELF-CONFIDENCE	TENACITY	DRIVE	INITIATIVE	PERSONAL INVOLVEMENT	ADAPTABILITY & FLEXIBILITY	LOYALTY/INTEGRITY/VALUES	RATING
GRADE												,24*	,20*							,22*
AGE	-,27**				-,23*			-,21*								-,26**				
LANGUAGE	-,31*						-,33*							-,34*						
SEX				,19*			,20*				,23*	,24*		,24*	,23*	,24*	,30**			,26**
RACE	-,25*						-,25**	-,19*						-,29**						
FLEXIBILITY																				
PERSONAL DEVELOPMENT	,26**	,25**	,22**		,22**			,23*	,23*	,19*	,21*	,24**	,36**		,21*	,27**				
ANALYTICAL SKILLS	,22*				,26**	,27**		,19*	,20*	,24*		,19*	,34**		,23*	,29**		,28**		
JUDGEMENT													,23*							,25**
DECISIVENESS					,20*									,20*				,26**		,21*
PLANNING & ORGANIZING	,23*	,21*	,19*	,24*	,23*	,20*			,29**	,23*	,19*	,22*	,26**	,21*	,26**	,23*	,26**	,31**		,33**
DELEGATION				,26**						,32**										,19*
CONTROL																				
INTERPERSONAL SENSITIVITY													,19*							
GROUP LEADERSHIP													,22*			,25**				
INDIVIDUAL LEADERSHIP	,22*		,20*						,36**		,24**	,21*	,33**	,19*		,30**	,27**	,31**		,22*
VERBAL COMMUNICATION	,32**		,19*			,22*	,20*		,20*		,27**	,31**	,32**	,23*	,21*	,33**	,31**			,19*
PRESENTATION SKILLS										,21*			,30**			,22*				,22*

*p < 0,05 **p < 0,01, (two-tailed)

Biographical data

Grade correlates significantly with Communication, Self-confidence and Rating. This confirms a general observation that the more senior participants performed better in the assessment centre. Age correlated negatively with four criteria, i.e. Analytical Skills, Planning and Organizing, Leadership and Initiative, indicating that younger participants did better on these dimensions.

Race and sex correlated significantly with a number of criteria, including Rating. These correlations should however be interpreted with circumspection since females and Blacks formed only a small part of the total sample of 110, i.e. females 11 and Blacks 10. The significant correlations could therefore be sample specific. Since reliable information about language was not always obtained, significant correlations with this variable should likewise be treated with care.

Assessment centre dimensions

Seven of the assessment centre dimensions correlated significantly with Rating, with Judgement and Planning and Organizing significant at the 1% level. The other significant predictors were Decisiveness, Delegation, Individual Leadership, Verbal Communication and Presentation Skills.

Furthermore, two predictors, i.e. Personal Development and Analytical Skills, correlated significantly with most criterion variables (Personal Development with 12 and Analytical Skills with 11 criterion variables). Both dimensions had correlations of ,18 with Rating, just outside the 5% level of significance (p = ,054 and ,053 respectively).

Intercorrelations between Groupings of Predictor and Criterion Variables

Classification of predictor and criterion variables into Personal Dimensions and Management Areas were presented in Tables 1 and 2 respectively.

Intercorrelations between groupings of predictor and criterion variables are shown in Table 5.

TABLE 5
INTERCORRELATIONS BETWEEN GROUPINGS OF PREDICTOR AND CRITERION VARIABLES (N = 110)

Criterion Variables Predictor variables	Personal dimensions	Decision making	Administrative skills	Leadership	Communication	Rating
Personal dimensions	0,26**	0,23*		0,25**	0,24**	0,26**
Decision-making	0,28**	0,21*	0,19*	0,21*	0,19*	0,26**
Administrative skills	0,19*	0,19*		0,24**		0,22*
Leadership	0,28**			0,22*		0,25**
Communication	0,34**	0,25**	0,23*	0,27**	0,34**	0,33**

* $p < 0,05$ ** $p < 0,01$, two-tailed.

Table 5 shows that Personal Dimensions and all four Management Areas, i.e. Decision-making, Administrative Skills, Leadership and Communication correlated significantly with Rating. All areas, except Administrative Skills, correlated significantly at the 1% level.

Furthermore, Decision-making and Communication correlated significantly with all criterion variables and Personal Dimensions with all criterion variables except Administrative Skills.

This table shows that when groups of variables are compound into comparable dimensions, significant positive correlations are found, pointing to validity in the whole system.

Multiple Regression Analysis

Stepwise multiple regression and simple regression analyses were computed for three sets of data, i.e. all predictor variables with Rating, assessment centre predictor variables with Rating and assessment centre Management Areas with Rating.

Results are shown in Table 6.

TABLE 6
MULTIPLE CORRELATIONS BETWEEN PREDICTOR VARIABLES AND RATING (N = 110)

PREDICTOR VARIABLES	Multiple R	F	Adjusted R Square	Multiple R after Shrinkage[a]
1. Predictor variables, including biographical data	0,54	2,34	0,16	0,41
2. Predictor variables without biographical data	0,49	2,31	0,14	0,37
3. Assessment centre management areas as predictor variables	0,40	3,92	0,12	0,34

Note: Planning and Organizing came out first in the stepwise multiple regression, irrespective of the inclusion of biographical data.

[a] The ratio between the number of variables and the number of subjects necessitated application of the shrinkage formula.

Results from Table 6 indicate that the addition of biographical data increased the validity of the prediction. Combination of variables into Management Areas resulted in a lower validity co-efficient. Both the above trends were expected and corroborate previous research.

Simple regression analysis

Simple regression analysis was also computed. Results are shown in Table 7.

It is clear that Planning and Organizing, Personal Development and Verbal Communication contribute most to the prediction of the total summed rating. Surprising, however, is the negative beta weights of Judgement and of Control. Because these variables load negatively in the prediction of Rating, they should be used with care when running future MMAC's. The content and measurement method should probably be scrutinized and adjusted first in order for these variables to load positively in the prediction of Rating.

TABLE 7
RESULTS OF SIMPLE REGRESSION ANALYSIS (N = 110)

PREDICTOR VARIABLES	BETA WEIGHTS
Age	0,011
Language	–0,084
Sex	0,189
Race	–0,079
Flexibility	0,060
Personal Development	0,198
Analytical Skills	0,031
Judgement	–0,175
Decisiveness	0,086
Planning and Organizing	0,277
Delegation	0,061
Control	–0,089
Interpersonal Sensitivity	–0,039
Group Leadership	–0,018
Individual Leadership	0,011
Verbal Communication	0,123
Presentation Skills	0,070

Canonical Correlations

The criterion in the multiple correlation, Rating, was calculated by simply adding dimension scores of the BARS. If an optimum combination of the criteria had been used, instead of just summing criterion scores, a higher correlation would be expected.

This was, in fact, achieved by applying the canonical correlation technique.

Three sets of canonical correlations were computed, as follows:

1. Predictor variables, including biographical data, with criterion variables.
2. Predictor variables, without biographical data, with criterion variables.
3. Assessment centre Management Areas with BARS Management Areas.

The results are shown in Table 8.

TABLE 8
CANONICAL CORRELATIONS BETWEEN PREDICTOR AND CRITERION VARIABLES (N = 110)

PREDICTOR VARIABLES	CANONICAL CORRELATION
1. Predictor variables including biographical data with criterion variables	0,70
2. Predictor variables without biographical data with criterion variables	0,67
3. Assessment Centre management areas with BARS management areas	0,43

Table 8 indicates that canonical correlations between predictor and criterion variables are substantially higher than the multiple correlations between the predictors and Rating for each of the three sets of data analyzed. This indicates that much is probably to be gained by using an optimum combination of weights in the criterion rather than the simple addition of the criterion dimension scores as was the case in this study.

DISCUSSION

From the results, it is clear that whether one takes simple correlations, multiple correlations or canonical correlations into account, all of these coefficients indicate the validity of this assessment centre.

These results are very positive if one considers the independant nature of the criteria, the fact that assessment centre dimensions and BARS were not directly comparable, and furthermore, that Rating was determined merely by simple addition of BARS scores and that no weighting took place.

In view of the above, a multiple correlation of 0,37 (after shrinkage) between assessment centre dimensions and Rating is indeed gratifying.

When biographical data are included the multiple correlation is slightly higher (0,41 after shrinkage). This finding is in line with some previous work that demonstrates the value of biographical information as predictors of performance (Schmitt, Gooding, Noe & Kirsch, 1984).

The slightly lower multiple correlation of 0,34 (after shrinkage) between groupings of dimensions as predictors (i.e. Management Areas) and Rating was expected, because of the loss of specific variability.

The significant correlation between grade and Rating, indicating that the more senior managers performed better in the assessment centre, was also expected. Negative correlations between age and some criterion variables (Analytical Skills, Planning and Organizing, Leadership and Initiative), signifying that younger managers did better than their older colleagues on these dimensions, corroborate previous findings (Thornton & Byham, 1982). Correlations between the other biographical variables, i.e. language, sex and race are probably sample specific and should be interpreted with caution.

The outstanding feature of correlations between assessment centre dimension scores and criterion ratings is the high number of significant correlations (9 of the 13 dimensions correlating significantly with either Rating or the majority of the BARS). This tendency is even more prevalent when predictor and criterion variables are combined to form respective sets of Personal Dimensions and four Management Areas, i.e. Decision-making, Administrative Skills, Leadership and Communication. All five of these predictors correlated significantly with Rating.

It may be appropriate to look at the assessment centre dimensions that did not correlate significantly with any criterion variable, i.e. Control, and Flexibility.

The inability of Control to predict performance on any of the criteria is probably due to restriction of range. Table 3 indicates that Control has the lowest mean score of all assessment centre dimensions, and inspection of the distribution of this variable shows that it is skewed to the right. This points towards either too strict evaluation or deficiencies in managers' ability to exercise control.

The insignificant correlations between Flexibility and criterion variables are in line with a finding in a study by Britz (1984) in which Flexibility did not predict any of five MBO criteria. Special care was taken with the measurement of Flexibility in the MMAC and two facets, i.e. mental and behavioural flexibility, were measured.

The negative results of Flexibility may be due partly to behaviour inconsistency and partly to measurement problems. For assessment centre behaviour to be considered as consistent, Guion (1987) put forward two requirements: firstly, it must be shown that the behaviour is indeed useful behaviour on the job and, secondly, the behaviour in the assessment centre should be consistent with behaviour shown on the job. The first requirement is probably met by the inclusion of Adaptability and Flexibility as one of the BARS and indicates that senior managers consider this dimension to be important for successful job performance. The second requirement, i.e. that behaviour shown in the assessment centre should be consistent with that shown on the job is probably not fully met. It may be that flexibility shown in the assessment centre just does not translate unchanged to the large number of diverse situations on the job calling for various forms of this behaviour. An alternative explanation is that flexibility is simply not properly measured in the assessment centre. It may for instance be that combining mental and behavioural measures of flexibility to form a more global measure reduced discriminative ability of this dimension.

In view of the valuable information provided by Flexibility for increasing self-insight and its consequent impact on self development, one would be hesitant to suggest its exclusion from an assessment centre for middle managers altogether. One would rather suggest that a good look is taken at the way in which Flexibility is measured, and more specifically the scoring procedure, whereby a "3", or average rating is awarded unless sufficient evidence is rendered to move the score away from a "3".

In view of recurring criticism in the literature that assessment centres are most predictive of advancement criteria – viz a viz performance criteria – (Klimoski & Strickland, 1977, 1981; Turnage & Muchinsky, 1984; Sackett & Ryan, 1985; Klimoski & Brickner, 1987) the validation results of the MMAC provides solid additional evidence of the ability of assessment centres to predict on the job performance.

It is indeed surprising that BARS which measure performance in behavioural terms have been used so seldom in assessment centre research. The only studies using BARS as performance criteria that could be traced are those by Worbois (1975, using behavioural scales), Huck (1974), Huck and Bray (1976), Slivinski (1977), Schmitt and Noe (1984), and Gaugler, Rosenthal, Thornton & Bentson (1987). In all these studies consistently positive results have been reported, although in studies where both BARS and potential or advancement criteria were used validity figures of the latter were higher.

It is equally surprising that with the exception of a recent – as yet unpublished – South African study (Britz, 1984), MBO results, as a performance measure, have not been used as criteria for assessment centre validation studies.

In the study by Britz which comprised a sample of over 600 middle and senior managers of the South African Transport Services (SATS) highly significant simple and multiple correlations between assessment centre ratings and various measures of MBO criteria were found (with multiple correlations between a total assessment centre score and the MBO criteria varying between 0,59 and 0,76).

The results of these two South African studies do indeed provide evidence that assessment centres, if constructed and implemented carefully, predict managerial *performance* whether it be in *behavioural* terms (BARS) or in terms of the achievement of *results* (MBO criteria).

Furthermore, the question arises why BARS and MBO types of criteria have not more often been used in validation studies. An obvious answer is that it is much easier to use advancement criteria because they are mostly readily available, whereas the construction of BARS for instance is time consuming and costly if used only for validation purposes.

The outcome of the emphasis on promotion and salary growth criteria is that useful information about assessment centre behaviour and work behaviour has not been forthcoming (Schmitt et al., 1984). It is not surprising therefore that even today evidence of construct validity of dimensions used in assessment centres is not encouraging (Klimoski & Brickner, 1987). It is small wonder that the same authors ask the question: why do assessment centres work?

What seems needed are well controlled validity studies in which multiple performance criteria, hard and soft, are researched. They could include performance criteria which form part of the ongoing management system, eg. MBO or Performance Management results, or multiple purpose criteria such as specially designed performance rating scales which can be used for measuring the effect of training and development efforts, for diagnosing performance problems or for periodic performance reviews, or other output measures applicable to the specific situation.

APPENDIX 1 – PART I
INTERCORRELATION MATRIX (N = 110)

	GRADE	AGE	LANGUAGE	SEX	RACE	FLEXIBILITY	PERSONAL DEVELOPMENT	ANALYTICAL SKILLS	JUDGEMENT	DECISIVENESS	PLANNING & ORGANIZING	DELEGATION	CONTROL	INTERPERSONAL SENSITIVITY	GROUP-LEADERSHIP	INDIVIDUAL LEADERSHIP	VERBAL COMMUNICATION	PRESENTATION SKILLS	ANALYTICAL SKILLS	JUDGEMENT	ORGANIZATIONAL & ENVIRONMENTAL SENSITIVITY	DECISIVENESS	PLANNING & ORGANIZING	DELEGATION	CONTROLLING	LEADERSHIP	INTERPERSONAL RELATIONS	NEGOTIATING SKILLS	DEVELOPMENT OF SUBORDINATES	COMMUNICATION	SELF-CONFIDENCE	TENACITY	DRIVE	INITIATIVE	PERSONAL INVOLVEMENT	ADAPTABILITY & FLEXIBILITY	LOYALTY/INTEGRITY/VALUES	RATING
GRADE	1,00																																				
AGE	,20*	1,00																																			
LANGUAGE	,32*		1,00																																		
SEX				1,00																																	
RACE		,20*	,49**		1,00																																
FLEXIBILITY						1,00																															
PERSONAL DEVELOPMENT	,23*	-,29**					1,00																														
ANALYTICAL SKILLS		-,26**					,37**	1,00																													
JUDGEMENT						,23*	,20*	,41**	1,00																													
DECISIVENESS								,28**	,24**	1,00																												
PLANNING & ORGANIZING		-,24*						,41**	,39**	,20**	1,00																											
DELEGATION	,26**						,20*	,31**	,35**		,48**	1,00																											
CONTROL									,26**		,39**	,34**	1,00																									
INTERPERSONAL SENSITIVITY	,24*					,47**	,19*	,29**	,28**	,26**				1,00																								
GROUP/LEADERSHIP		-,21*				,33**	,25**	,30**	,19*	,24*		,23*		,29**	1,00																							
INDIVIDUAL LEADERSHIP				,27**		,46**	,37**	,48**	,24*	,28**	,28**			,48**	,42**	1,00																						
VERBAL COMMUNICATION				,29**		,22*		,28**		,25**	,19*			,27**	,35**	,51**	1,00																					
PRESENTATION SKILLS							,23*	,44**	,28**	,21*		,23*		,23*	,44**	,27**	,21*	1,00																				
ANALYTICAL SKILLS	-,27**	-,31*		-,25*		,26**	,22*			,23*						,22*	,32**		1,00																			
JUDGEMENT							,25**				,21*								,64**	1,00																	
ORGANIZATIONAL & ENVIRONMENTAL SENSITIVITY						,22*					,19*					,20*	,19*		,50**	,51**	1,00																
DECISIVENESS				,19*							,24*	,26**							,42**	,51**	,45**	1,00															
PLANNING & ORGANIZING		-,23*					,22*	,26**		,20*	,23*								,54**	,49**	,34**	,33**	1,00														
DELEGATION								,27**			,20*						,22*		,55**	,45**	,43**	,36**	,65**	1,00													
CONTROLLING		-,33*	,20*	-,25**													,20*		,56**	,49**	,43**	,49**	,54**	,52**	1,00												
LEADERSHIP		-,21*			-,19*		,23*	,19*											,48**	,56**	,51**	,43**	,57**	,50**	,57**	1,00											
INTERPERSONAL RELATIONS							,23*	,20*			,29**					,36**	,20*		,33**	,38**	,49**	,19*	,32**	,31**	,35**	,46**	1,00									
NEGOTIATING SKILLS							,19*	,24*			,23*	,32**						,21*	,45**	,40**	,53**	,44**	,34**	,29**	,44**	,40**	,52**	1,00									
DEVELOPMENT OF SUBORDINATES				,23*			,21*				,19*					,24**	,27**		,33**	,39**	,44**	,34**	,51**	,54**	,55**	,57**	,37**	,38**	1,00							
COMMUNICATION	,24*			,24*			,24**	,19*			,22*					,21*	,31**		,44**	,39**	,43**	,30**	,39**	,41**	,35**	,45**	,39**	,45**	,65**	1,00						
SELF-CONFIDENCE	,20*						,36**	,34**	,23*	,20*	,26**			,19*	,22*	,33**	,32**	,30**	,49**	,45**	,49**	,35**	,32**	,40**	,36**	,42**	,47**	,62**	,39**	,44**	1,00						
TENACITY			-,34*	,24*	-,29**						,21*					,19	,23*		,46**	,60**	,49**	,58**	,30**	,30**	,52**	,49**	,29**	,35**	,48**	,62**	,62**	1,00					
DRIVE				,23*			,21*	,23*			,26**						,21*		,49**	,54**	,57**	,52**	,42**	,40**	,55**	,62**	,45**	,55**	,56**	,71**	,53**	,75**	1,00				
INITIATIVE		-,26**		,24*			,27**	,29**			,23*				,25**	,30**	,33**	,22*	,64**	,58**	,59**	,45**	,49**	,53**	,50**	,54**	,45**	,56**	,46**	,48**	,44**	,60**	,56**	1,00			
PERSONAL INVOLVEMENT				,30**							,26**					,27**	,31**		,30**	,31**	,46**	,36**	,21*	,26**	,25**	,44**	,43**	,37**	,27**	,41**		,27**	,38**	,29**	1,00		
ADAPTABILITY & FLEXIBILITY								,28**		,26**	,31**					,31**			,27**	,31**	,41**		,27**	,29**		,21*	,60**	,41**	,23*	,22*	,42**	,51**	,38**	,40**	,31**	1,00	
LOYALTY/INTEGRITY/ VALUES																			,24*	,48**	,32**	,35**	,33**	,25**	,37**	,49**	,38**	,34**	,67**	,73**	,67**	,81**	,84**	,61**	,47**	,57**	1,00	
RATING	,22*			,26**					,25**	,21*	,33**	,19*				,22*	,19*	,22*	,71**	,73**	,72**	,62**	,66**	,66**	,70**	,75**	,63**	,69**	,68**	,67**	,73**	,67**	,81**	,84**	,61**	,47**	,57**	1,00

REFERENCES

Britz, P.J. (1984). *Die validering van 'n bestuursbeoordelingsentrum.* Unpublished doctoral dissertation, University of Pretoria, Pretoria.

Gaugler, B.B., Rosenthal, D.B., Thornton, G.C., & Bentson, C. (1987). Meta-analysis of assessement centre validity. Journal of Applied Psychology, 72, (3), 493-511.

Guion, R.M. (1987). Changing views for personnel selection research. *Personnel Psychology, 40,* 199-213.

Huck, J.R. (1974). *Determinants of assessment centre ratings for white and black females and the relationships of these dimensions to subsequent performance effectiveness.* Unpublished doctoral dissertation, Wayne State University.

Huck, J.R., & Bray, D.W. (1976). Management assessment centre evaluations and subsequent job performance for Black and White females, *Personnel Psychology, 29,* 13-30.

Klimoski, R.J., & Strickland, W.J. (1977). Assessment centres: valid or merely prescient. *Personnel Psychology, 30,* 353-363.

Sackett, P.R., & Ryan, A.M. (1985). A review of recent assessment centre research, *Journal of Management Development, 4,* 13-27.

Schmitt, N., Gooding, R.Z., Noe, R.A., & Kirsch, M. (1984). Meta-analysis of validity studies published between 1964 and 1982 and study characteristics. *Personnel Psychology, 37,* 407-422.

Slivinski, L.W., Grant, K.W., Bourgeois, R.P., & Pederson, L.D. (1977). *Development and application of a first level management assessment centre.* Ottawa, Ontario, Canada: Public Service Commission of Canada, Personnel Psychology Centre.

Slivinski, L.W., Bourgeois, R.P., Pederson, L.D., Mc Closkey, J.L., McDonald, V.S., & Crooks, L.A. (1981). *Identification of senior executive potential. Development and application of an assessment centre.* Ottawa, Ontario, Canada: Public Service Commission of Canada, Personnel Psychology Centre.

Smith, P.C., & Kendall, L.M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology, 47,* 140-155.

Spangenberg, H.H. (in press). *Constructing, implementing and validating an assessment centre.*

Spangenberg, H.H, Esterhuyse, J.J., Visser, J.H., Briedenhann, J.E. & Calitz, C.J. (1989). *Construction of BARS for the measurement of managerial performance. Journal of Industrial Psychology, 15*(1), 22-27.

Thornton, G.C., & Byham, W.C. (1982). *Assessment centres and managerial performance.* New York: Academic Press.

Turnage, J.J., & Muchinsky, P.M. (1984). A comparison of the predictive validity of assessment centre evaluations versus traditional measures in forecasting supervisory job performance: interpretive implications of criterion distortion for the assessment paradigm. *Journal of Applied Psychology, 69,* 207-213.

Worbois, G.M. (1975). Validation of externally developed assessment procedures for identification of supervisory potential. *Personnel Psychology, 28,* 77-91.