

**THE ASSESSMENT CENTRE AS A PREDICTOR OF  
MANAGERIAL TRAINING RESULTS**

by

**BENJAMIN HUGH OLIVIER**

submitted in part fulfilment of the requirements  
for the degree of

**MASTER OF ADMINISTRATION**

in the subject

**INDUSTRIAL PSYCHOLOGY**

at the

**UNIVERSITY OF SOUTH AFRICA**

**SUPERVISOR: PROF H J KRIEK**

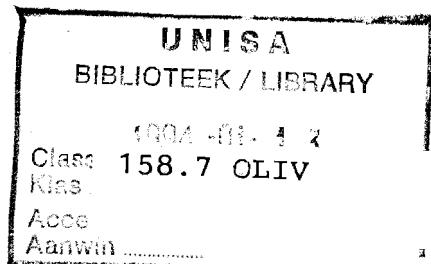
**MAY 1994**

## ACKNOWLEDGEMENTS

I would like to express my thanks and appreciation to the following people:

- \* Prof H. Kriek for his guidance and support.
- \* Mrs Gerda Coetzee and Mrs Daleen Niesen for typing parts of this dissertation.
- \* Mr Trevor Amos and Mrs Faris for the proof reading and language refinement of the contents.
- \* The SA Army for making the necessary data available to me.
- \* My wife who had to keep the children entertained and busy while I was isolated in my study for such long periods of time.

*This dissertation is dedicated to my mother who passed away on the 27th of January 1994. I know that she is experiencing the peace and joy she had waited so long for; the peace and joy which she so richly deserves.*



01556782

## SUMMARY

THE ASSESSMENT CENTRE AS  
A PREDICTOR OF MANAGERIAL  
TRAINING RESULTSby  
Benjamin Hugh Olivier

Supervisor : Prof H Kriek  
Degree : M Admin  
Subject : Industrial Psychology

The aim of this study is to investigate the validity of an assessment centre as a predictor of managerial training results. For this purpose, the assessment ratings of 77 white, male, middle managers, who had taken part in a Middle Management Assessment Centre (MMAC) in a South African service organisation, were compared to the results they achieved on a Senior Management Training Course.

From the statistical analyses conducted, correlation coefficients, corrected for restriction of range, of -0,03, 0,32 and 0,22 were obtained between the Overall Assessment Rating and a theoretical, practical and total grade criterion respectively. These results indicated that the MMAC was a valid predictor of practical and total course grades, but an invalid predictor of theoretical course grades.

Explanations for the results obtained are provided and three recommendations are made concerning the use of the MMAC to predict training results. Directions for possible future research are also provided.

**Key terms.** Assessment centres; Prediction of training results; Managerial training; Assessment centre validity; Training performance; Service organisations; Simulation exercises; Managerial performance; Managers; Selection for training.

## OPSOMMING

### DIE TAKSEERSENTRUM AS 'N VOORSPELLER VAN BESTUURSOPLEIDING RESULTATE deur Benjamin Hugh Olivier

Studieleier : Prof H Kriek  
Graad : M Admin  
Vak : Bedryfsielkunde

Die doel van hierdie studie is om die geldigheid van 'n takseersentrum as 'n voorspeller van bestuursopleiding resultate te ondersoek. Vir hierdie doel is die takseeraanslae van 77 blanke, manlike bestuurders, wat aan 'n takseersentrum vir middelbestuur in 'n Suid-Afrikaanse diensorganisasie deelgeneem het, met die resultate wat hulle tydens 'n Senior Bestuur Opleidingskursus behaal het, vergelyk.

Die statistiese ontleding wat uitgevoer is het korrelasiekoeffisiënte, na regstelling vir die beperking van verspreiding (correction for restriction of range), van -0,03, 0,32 en 0,22 tussen die Totale Punt Beoordeling van die takseersentrum en drie opleidingskriteria opgelewer. Hierdie drie opleidingskriteria was 'n teoretiese-, praktiese-, en totale opleidingspunt onderskeidelik.

Verduidelikkings vir die resultate wat behaal is word verskaf en drie aanbevelings word gemaak met betrekking tot die gebruik van die takseersentrum om opleidingsresultate te voorspel. Rigtigs vir moontlike toekomstige navorsing word ook aangedui.

## TABLE OF CONTENTS

	Page
<b>SUMMARY</b>	<b>(iii)</b>
<b>LIST OF TABLES</b>	<b>(xii)</b>
<b>LIST OF FIGURES</b>	<b>(xv)</b>
<b>APPENDICES</b>	<b>(xvi)</b>
<b>CHAPTER 1:INTRODUCTION</b>	<b>1</b>
<b>1.1 BACKGROUND TO THE STUDY</b>	<b>1</b>
1.1.1 The importance of management to the success of an organisation	1
1.1.2 The importance of management to a larger society	2
1.1.3 Necessity of training and developing managers	3
<b>1.2 RATIONALE FOR THE STUDY</b>	<b>4</b>
1.2.1 The necessity of predicting managerial training performance	5
(a) Return on training investment	5
(b) Ensuring sufficiently trained personnel by eliminating training failures	6
(c) Improving training effectiveness	7
1.2.2 The necessity for training selection to be valid	7

	Page
<b>1.3 AIM OF THE STUDY</b>	<b>8</b>
1.3.1 The assessment centre as a predictor of success	8
1.3.2 The necessity of validating an assessment centre used to predict training success	9
1.3.3 General aim	10
1.3.4 Specific objectives	11
1.3.5 The research question	11
<b>1.4 SCOPE OF THE REPORT</b>	<b>11</b>
1.4.1 General orientation	11
1.4.2 Specific chapters	12
<b>1.5 SUMMARY</b>	<b>12</b>
<b>CHAPTER 2:THE PREDICTION OF MANAGERIAL TRAINING RESULTS</b>	<b>14</b>
<b>2.1 INTRODUCTION</b>	<b>14</b>
<b>2.2 THE CONTENT OF CURRENT-DAY MANAGERIAL TRAINING PROGRAMMES</b>	<b>14</b>
<b>2.3 THE USE OF TRAINING RESULTS AS A CRITERION OF TRAINING SUCCESS</b>	<b>14</b>
2.3.1 The measurement of training results	15
2.3.2 Learning achieved as a measure of training success	16
2.3.3 Methods of quantifying learning achieved	16
(a) Training instructor judgments	17
(b) Scores on paper-and-pencil knowledge tests	18
(c) Performance test scores	19
2.3.4 Summary	20

	Page
<b>2.4 TRAINING PREDICTORS AND THEIR VALIDITY</b>	<b>20</b>
<b>2.4.1 Approaches to prediction</b>	<b>20</b>
<b>2.4.2 Validity of instruments used to predict managerial training results</b>	<b>21</b>
(b) Background information	22
(c) Interviews	24
(d) Personality inventories	26
(e) Peer assessment	27
(f) Psychometric ability tests	29
(g) Work-sample performance tests	32
(h) Situational work-sample performance tests	36
<b>2.5 SITUATIONAL TESTS AND THE ASSESSMENT CENTRE METHOD</b>	<b>41</b>
<b>2.6 SUMMARY</b>	<b>42</b>
<b>CHAPTER 3:THE ASSESSMENT CENTRE METHOD</b>	<b>43</b>
<b>3.1 INTRODUCTION</b>	<b>43</b>
<b>3.2 DESCRIPTION OF THE ASSESSMENT CENTRE METHOD</b>	<b>43</b>
<b>3.2.1 Assessment centres defined</b>	<b>43</b>
<b>3.2.2 The assessment centre approach versus psychometric prediction</b>	<b>44</b>
<b>3.2.3 Underlying rationale of assessment centres</b>	<b>44</b>
<b>3.2.4 Instruments and exercises used in assessment centres</b>	<b>45</b>
<b>3.2.5 Industrial applications of assessment centres</b>	<b>46</b>

<b>3.3</b>	<b>THE ASSESSMENT CENTRE AS A PREDICTOR OF MANAGERIAL PERFORMANCE</b>	<b>47</b>
3.3.1	Criteria utilised in assessment centre validity studies	48
3.3.2	Assessment centre measures used as predictors	49
3.3.3	Predictive validity research results	50
3.3.4	Discussion	52
<b>3.4</b>	<b>SUMMARY</b>	<b>53</b>
<b>CHAPTER 4:THE ASSESSMENT CENTRE AS A PREDICTOR OF MANAGERIAL TRAINING RESULTS</b>		<b>55</b>
<b>4.1</b>	<b>INTRODUCTION</b>	<b>55</b>
<b>4.2</b>	<b>STUDIES IN WHICH MANAGERIAL TRAINING RESULTS WERE UTILISED TO VALIDATE ASSESSMENT CENTRES</b>	<b>55</b>
4.2.1	Study 1	56
4.2.2	Study 2	57
4.2.3	Study 3	57
4.2.4	Study 4	58
4.2.5	Study 5	61
4.2.6	Study 6 and 7	63
<b>4.3</b>	<b>DISCUSSION</b>	<b>65</b>
<b>4.4</b>	<b>SUMMARY</b>	<b>67</b>
<b>CHAPTER 5:METHOD OF INVESTIGATION</b>		<b>69</b>
<b>5.1</b>	<b>INTRODUCTION</b>	<b>69</b>
<b>5.2</b>	<b>HYPOTHESIS FORMULATION</b>	<b>69</b>

<b>5.3</b>	<b>SETTING</b>	<b>69</b>
5.3.1	Purpose of the Middle Management Assessment Centre (MMAC)	70
5.3.2	Participation	70
5.3.3	Candidate acceptance for the Senior Management Training Course (SMTTC)	70
<b>5.4</b>	<b>SAMPLE</b>	<b>71</b>
<b>5.5</b>	<b>MEASURING INSTRUMENTS</b>	<b>73</b>
5.5.1	The MMAC	73
(a)	Development of the MMAC	73
(b)	Assessment staff	74
(c)	Assessment process	75
5.5.2	Independent variables	76
(a)	Consensus dimension ratings	76
(b)	Average dimension score	77
(c)	Weighted Overall Assessment Rating (OAR)	79
(d)	Moderator variables	80
(e)	Summary of independent variables	82
5.5.3	Dependent variables	82
(a)	The use of training grades as criteria of success	82
(b)	Multiple versus composite criteria	84
(c)	Measurement of multiple criteria used	85
(d)	Measurement of the composite criterion used	86
<b>5.6</b>	<b>RESEARCH DESIGN AND PROCEDURE</b>	<b>86</b>
5.6.1	Assessment centres and predictive validity	86
5.6.2	Obtaining predictor and criterion data	87

	Page
<b>5.7 DATA-ANALYSIS METHODS</b>	<b>87</b>
5.7.1 Intercorrelations between individual dimension ratings, average dimension rating and OAR	88
5.7.2 Intercorrelations between training criteria scores	89
5.7.3 Validity of predicting measures with training results as the criterion of success	89
5.7.4 The contribution of moderator variables to assessment centre validity	90
5.7.5 Multiple linear regression analysis	91
5.7.6 Multiple stepwise regression analysis	92
5.7.7 Analysis and interpretation of results	93
<b>5.8 SUMMARY</b>	<b>93</b>
<b>CHAPTER 6:RESEARCH RESULTS</b>	<b>95</b>
6.1 INTRODUCTION	95
6.2 INTERCORRELATIONS BETWEEN VARIABLES	95
6.2.1 Intercorrelations between predictors	95
6.2.2 Intercorrelations between criteria	97
6.3 CORRELATIONS BETWEEN PREDICTING MEASURES AND TRAINING CRITERIA	98
6.3.1 Validity of individual dimension ratings	99
6.3.2 Validity of average dimension rating	100
6.3.3 Validity of the weighted OAR	101
6.4 CONTRIBUTION OF MODERATOR VARIABLES TO ASSESSMENT CENTRE VALIDITY	101

	Page
<b>6.5 INDIVIDUAL AND COMBINED ACCURACY OF VARIOUS MMAC PREDICTOR COMBINATIONS IN PREDICTING TRAINING RESULTS</b>	<b>102</b>
6.5.1 Results with Theoretical Grade as the dependent variable	102
6.5.2 Results with Practical Grade as the dependent variable	103
6.5.3 Results with Total Grade as the dependent variable	103
<b>6.6 CONTRIBUTION OF OAR, AGE AND EXPERIENCE TO MMAC VALIDITY</b>	<b>103</b>
6.6.1 Unforced stepwise regression analysis	105
6.6.2 Forced stepwise regression analysis	107
<b>6.7 SUMMARY</b>	<b>108</b>
<b>CHAPTER 7:DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS</b>	<b>110</b>
<b>7.1 INTRODUCTION</b>	<b>110</b>
<b>7.2 DISCUSSION OF RESEARCH RESULTS</b>	<b>110</b>
7.2.1 Predictive validity of the MMAC	110
(a) Lack of point-to-point correspondence between predictor and criterion	111
(b) Utilising an assessment centre for which it was not designed	113
(c) Restriction of range	114
7.2.2 Moderators of MMAC validity	116
7.2.3 Accuracy of MMAC predictor combinations	117
7.2.4 Validity of weights used to calculate the OAR	117

<b>7.3</b>	<b>LIMITATIONS OF THE STUDY</b>	<b>118</b>
7.3.1	Size of the research sample used	118
7.3.2	Reliability of predictors and criteria	119
<b>7.4</b>	<b>CONCLUSIONS</b>	<b>119</b>
7.4.1	The assessment centre is a predictor of managerial training results	120
7.4.2	A combination of predictors is more valid than individual predictors	120
7.4.3	Levels of validity for predictors depend on the nature of the criteria used	121
7.4.4	The use of weights to calculate the final OAR does not improve the ability of the MMAC to predict training results	121
<b>7.5</b>	<b>RECOMMENDATIONS</b>	<b>121</b>
7.5.1	Trainees on the SMTC should be assessed according to behavioural dimensions	121
7.5.2	Dimensions representing successful training behaviour should be determined	122
7.5.3	Weights should not be used to calculate an OAR	122
<b>7.6</b>	<b>DIRECTIONS FOR FUTURE RESEARCH</b>	<b>122</b>
7.6.1	Utility of the assessment centre as a selection device for training programmes	122
7.6.2	Validating the MMAC with on-the-job performance	123
<b>7.7</b>	<b>SUMMARY</b>	<b>123</b>
<b>REFERENCE LIST</b>		<b>125</b>

**LIST OF TABLES**

<b>Table 2.1</b>	Levels of training evaluation, type of training criteria, and purpose of each type of criteria	16
<b>Table 2.2</b>	Validities of background information measures with training success as the criterion	24
<b>Table 2.3</b>	Validities of peer nominations with training performance as the criterion	28
<b>Table 2.4</b>	Validities of ability tests used to predict training results in managerial occupations	30
<b>Table 2.5</b>	Correlations between mental ability tests and training performance	30
<b>Table 2.6</b>	Correlations of work-sample and intelligence tests with training performance as the criterion	34
<b>Table 2.7</b>	Validity coefficients for different types of work-sample tests with training performance as the criterion	35
<b>Table 2.8</b>	Validity coefficients of situational tests with Police Academy training results as criteria .	38
<b>Table 2.9</b>	Correlations between situational tests and performance in officer training	39

	Page
Table 2.10 Correlations between situational tests and performance in military recruiter training	39
Table 3.1 Frequency of use of assessment centre instruments	47
Table 3.2 Categories of criteria utilised in assessment centre validity studies	48
Table 4.1 Assessment centre validity studies using training results as criteria	56
Table 4.2 Validity of assessment centre predictors with training grades as criteria	59
Table 4.3 Validity of traditional selection predic- tors compared to assessment centre predic- tors with training results as criteria	60
Table 4.4 Validity of various assessment centre predictors with training and job performance as criteria	62
Table 4.5 Correlations of extended interview (assessment centre) measures with training criteria	64
Table 5.1 Characteristics of the research sample	72
Table 5.2 Dimensions assessed in each situational exercise	75
Table 5.3 Calculation of the MMAC weighted OAR	80
Table 5.4 Summary of independent variables	83

Table 6.1	Matrix of intercorrelations between individual dimension ratings, average dimension rating and OAR	96
Table 6.2	Matrix of intercorrelations between theoretical, practical, and total training grades (criteria)	98
Table 6.3	Correlations between assessment centre predictors and training results (criteria)	99
Table 6.4	Validity coefficients of functional exposure, age and experience with training results as criteria	102
Table 6.5	Contribution of various predictor combinations to assessment centre validity	104
Table 6.6	Stepwise regression analysis : OAR, age and experience (unforced)	106
Table 6.7	Stepwise regression analysis : OAR, age and experience (all forced)	107
Table 7.1	Comparison of MMAC validity to results of meta-analysis of assessment centre validity conducted by Gaugler et al (1987)	111

**LIST OF FIGURES**

<b>Figure 2.1 Instruments used to predict managerial training results</b>	<b>22</b>
<b>Figure 5.1 Categories of MMAC dimensions</b>	<b>74</b>
<b>Figure 5.2 Ten-point rating scale used to rate dimensions during each situational exercise</b>	<b>78</b>

**APPENDICES**

Appendix A : Steps followed in the development of the MMAC	138
Appendix B : Example of a MMAC exercise-dimension matrix	139

## CHAPTER 1

### INTRODUCTION

#### 1.1 BACKGROUND TO THE STUDY

##### 1.1.1 The importance of management to the success of an organisation

Katz and Kahn (1978), Kroon (1990) and Schein (1980), state that since the beginning of time organisations have been formed to achieve objectives not possible through individual effort. According to Kroon, as these organisations grew and became more sophisticated in their functioning, the role of management became indispensable as a coordinating agent and provider of direction. For this reason, Kroon (1990) and Walker (1992) refer to management as one of the most important human activities in organisations today.

This is supported by the view of Byham and Thornton (1970), who state that the manager holds an important job in any organisation, as he must plan, coordinate, and control the execution of work. Besides providing a job situation which is motivating and stimulating to his employees, Byham and Thornton envisage him as the one who must give new direction to the organisation in the dynamic world of business, and to do this, the authors claim that he must have a broad perspective of the internal and external systems which affect organisation effectiveness. Drucker (1968) agrees with this view, calling the manager the dynamic, life-giving element in every business.

In their study examining differences in effectiveness among companies competing in the same market, Peters and Waterman (1982) found that management played an important role in the success of an organisation. They found that most companies within a given industry, having roughly the same potential access to resources, differed significantly in their operations and results. While potential access to resources may be roughly equal, actual access was found to differ markedly. By carefully analysing the successful organisations, Peters and Waterman found that it was differences in the management of these organisations which allowed them to acquire more than their share of the desired resources, while others fared less well.

Camp, Blanchard and Huszczo (1986) claim that because managers are responsible for the planning, organising and controlling of activities within organisations, they are ultimately responsible for the effectiveness of their organisation. Drucker (1968), takes this viewpoint further, saying that if objectives are not achieved, management is to blame. According to Camp et al, managers must ensure that employees have the knowledge and skills required to perform their jobs successfully and also to keep employees motivated. The authors consequently argue that if employees do not achieve their level of competence, it is ultimately considered to be management's fault.

#### **1.1.2 The importance of management to a larger society**

Human and Human (1989), in a study on the importance of management in a future South Africa, stress the fact that management, as a distinct occupational group with specific tasks to perform and specific responsibilities to meet, is becoming an increasingly important force in the current and future development

of a society. According to the authors, the successful development of a society would appear to be dependent to a high degree upon the effectiveness and efficiency of organisations in relation to the production and distribution of goods and services. The organisation is viewed by Human and Human (1989) as a developmental instrument which is directed and controlled by management, with its effectiveness and efficiency being largely a function of the quality of management. The authors feel that this assertion holds true for all types of organisations, and that management is important not only to the ensuring of the effectiveness and efficiency of their own organisations, but also to the creation of a better society at large.

#### **1.1.3 Necessity of training and developing managers**

In a discussion of human resource strategy, Walker (1992) states that as organisations grow and change, the need for high-quality management increases, with the emphasis on managing new products and services, technological advances, global business activities and alliances and improvement of competitive performance all creating strains on managerial abilities and adaptability. Walker states that the more rapid rate of change experienced in many organisations today, compared with a generation ago, compels special attention to the development of managerial talent.

For Kroon (1990), and McGehee and Thayer (1961), managers must be trained to a level of skill and knowledge so that their job performance contributes to the effectiveness and efficiency of their organisations. McGehee and Thayer emphasise that management at the middle and upper level of an organisation can, by a wrong decision, cost a company

more in a few minutes than any loss growing out of ineffective training of personnel at lower levels. Although training is no panacea for poor decision making, the authors believe it is possible that a considerable number of faulty decisions come from managers who are inadequately trained for the performance of their duties.

Byham and Thornton (1970) expand on this idea, stating that changes in the environment require a manager to learn new ways of managing an organisation, while differences between levels of management necessitate the acquisition of new knowledge and skills in order to succeed at a higher level of management. The authors therefore conclude that training and developing managers to manage effectively and efficiently both now and in the future would become a crucial issue for any organisation. This is confirmed by research conducted by Saari, Johnson, McLaughlin and Zimmerle (1988), who found that there appears to be an increased recognition of the efficacy of management training and education as strategies to improve organisational effectiveness and competitiveness. Their research shows that management training and education in the next five years will increase substantially, the main reasons being to update managers on changing concepts and skills.

## 1.2 RATIONALE FOR THE STUDY

As managers hold an important position in organisations today (Byham & Thornton, 1970; Camp et al, 1986; Drucker, 1968; Kroon, 1990; Peters & Waterman, 1982; Stewart & Stewart, 1976), and training and developing managers is critical to the growth and survival of organisations (Byham & Thornton, 1970; Kroon, 1990; Saari et al, 1988; Walker, 1992), the

decision as to who should attend managerial training programmes becomes crucial (Cascio, 1987; Noe & Schmitt, 1986; Russel, 1984). According to Gatewood and Feild (1990), this decision entails selecting applicants who have a high possibility of being successful, and to do this it is necessary to predict future managerial training performance.

#### 1.2.1 The necessity of predicting managerial training performance

From the available literature reviewed below, four main reasons for the importance of predicting managerial training performance can be identified.

##### (a) Return on training investment

According to Burke and Day (1986), managerial training and development programmes are implemented in most private and public organisations every year. Numerous studies have found that organisations devote a considerable amount of time, money and effort to training activities (Camp et al, 1986; Cascio, 1987; Goldstein, 1980; Russel, 1984), while Campbell, Dunnette, Lawler and Weick (1970), have documented a large volume of training activity which takes place in organisations, showing that almost all reasonably large corporations have formal, in-house training programmes. Concerning the financing of training activities, Campbell (1971) found that many corporate training budgets are comparable to those of many large universities.

Given this investment which organisations make in training, in terms of time, money and effort, Cascio (1987) found that a return on this training investment is expected. For Cascio, this return on investment is realised to a degree if all the trainees

are successful at mastering the training material and ultimately pass the training course. A way of ensuring a return on managerial training investment, propagated by Noe and Schmitt (1986) as well as by Taylor (1982), is to eliminate the possibility of early course drop-outs by selecting the right candidates to attend managerial training programmes.

(b) Ensuring sufficiently trained personnel by eliminating training failures

According to Camp et al (1986), a training needs assessment will identify the essential knowledge, skills and abilities (KSA) that a trainee needs to acquire on a training programme, but he will not perform, nor is he trainable, if he lacks the ability to develop the KSA in question. In discussing the ability of a person to be trained, Hinricks (1976) states that it is necessary to determine whether applicants possess the necessary abilities to learn and perform before being granted entrance to a training programme. Cascio (1987) supports this idea by stating that if candidates who are not able to master the content of managerial training courses are selected for training, an organisation may find itself with insufficient numbers of appropriately trained and qualified managers with which to cope with future needs and to deal with change. To prevent this becoming a reality, Camp et al (1986) suggest that an assessment be made of aspirant trainees in order to determine whether they are capable of developing the knowledge, skills and abilities being taught on the training course. The required pass-rate can then be achieved if unsuitable candidates are not granted entry to such a training course.

Downs (1970), Noe and Schmitt (1986), and Taylor (1982) use the term *trainability* when referring to

the ability of candidates to be trained, while Gordon and Kleiman (1976), and Camp et al (1986), use the term *prediction of trainability* to refer to the actual assessment of prospective trainees, to determine their trainability. Downs (1970), Noe and Schmitt (1986), Robertson and Downs (1979), Taylor (1982), and Taylor and Tajen (1948) claim that if trainability is predicted successfully, the right candidates can be selected to attend training programmes. The possibility of early course drop-outs and training failures will then be eliminated, and organisations will not find themselves with insufficient numbers of appropriately trained and qualified personnel with which to face the future.

(c) **Improving training effectiveness**

McGehee and Thayer (1961) and Taylor and Tajen (1948) claim that the advantages of selecting trainees in terms of probable training success is that it permits the concentration of training effort on those most likely to profit by it, increases the proficiency of the ultimate product, and ensures that training efforts are effective. According to Gordon and Cohen (1973), cost savings have resulted from the identification of individuals who will perform best in training, complete training most rapidly, or who can be permitted to omit portions of the training.

**1.2.2 The necessity for training selection to be valid**

According to Camp et al (1986), the decision as to who should attend a training programme implies that selected trainees will receive further training or education of some sort, either by means of a university-based programme or by means of internal management training courses. Russel (1984) states that in such cases training then becomes a key factor in determining promotions and career opportunities.

Quoting from fair employment court cases in the field of training, Russel claims that selection decisions for training are considered employment decisions, and are subject to fair employment legislation.

Cascio (1987) states that employment decisions have to be proven to be valid to comply with legislation, guidelines and principles governing the use of personnel selection procedures. Based on Gateway and Feild's (1990) view of selection validity, selection for training decisions can be proved valid if the selection device used effectively predicts training results.

### **1.3 AIM OF THE STUDY**

#### **1.3.1 The assessment centre as a predictor of managerial success**

The assessment centre has been used successfully for various purposes, such as predicting success in management jobs (Byham, 1970; Howard, 1974; Spangenberg, Esterhuyse, Visser, Briedenham & Calitz, 1989) and progress in management (Bray, 1964; Bray & Grant, 1966). The assessment centre method has also been praised as an important improvement over paper-and-pencil testing (Bray & Grant, 1966; Mitchel, 1975). Other assessment centres have even been developed for predicting women's advancement into middle management (Ritchi & Moses, 1983), Black leadership potential in South Africa (Charoux, 1990), and training performance in the military (Borman, 1982; Tziner & Dolan 1982).

The assessment centre literature also contains examples of assessment centres which have been used to select applicants for specific jobs. Bray and Campbell (1968) used an assessment centre to select salesmen and obtained a correlation of 0,51 between

performance on the assessment centre and ratings of field performance. Pynes and Bernardin (1989), reporting on a study to determine the predictive validity of an entry-level Police officer assessment centre, obtained a correlation of 0,20 between assessment centre performance and on-the-job performance.

#### **1.3.2 The necessity of validating an assessment centre used to predict training success**

If an assessment centre is used as a selection technique, Thornton and Byham (1982) state that its effectiveness must be demonstrated by evidence of predictive validity. This view is supported by Sackett (1987) who states that centres being used to select individuals for entry into a training programme are being used as aptitude tests to predict future training performance, and not as measures of current performance, and the appropriate validation model is criterion-related validity. In this regard, the article by Russel (1984) would appear to be relevant. In his article, Russel, after having reviewed various fair employment court cases in the field of training as well as fair employment legislation currently in force in the United States, concluded that employers have three strategies for validating selection for training techniques:

- \* They can validate training performance with job performance;
- \* They can validate the selection technique with job performance. However, if training performance is required for future employment opportunities, the training programme itself would have to be validated with job performance; or

- \* They can validate the selection technique with training performance.

Whichever validation strategy is followed, the Society for Industrial Psychology (1992), hereafter referred to as The Guidelines (1992), states that the selection procedure should be documented with appropriate evidence built on a foundation of systematic validation procedures such as those discussed in The Guidelines (1992).

### 1.3.3 General aim

At present a middle management assessment centre (MMAC) is being used by a service organisation as a screening device to select candidates for entry into an extensive senior management training course (SMTc). The MMAC was developed to select applicants who have the ability to master the content of the SMTc in order to ensure a return on training investment, and to eliminate course failures so that the organisation would be ensured of sufficiently trained managers to cope with the demands of the environment. The decision to utilise an assessment centre was based on existing research which supports the predictive validity of the assessment centre method (Gaugler, Rosenthal, Thornton & Bentson, 1987; Schmitt, Gooding, Noe & Kirsh, 1984). However, due to the fact that the MMAC is a labour intensive selection device requiring a substantial financial investment, it became necessary for the organisation to ensure that the assessment centre was doing what it was developed to do, namely, predict training performance. The need for a validity study with training performance as the criterion was thus identified, and a validity study was commissioned at the beginnig of 1992. The aim of this study is thus *to determine whether the MMAC predicts training results achieved on the SMTc.*

#### **1.3.4 Specific objectives**

The specific objectives of this study are the following:

- \* To review the literature on the prediction of training results, with emphasis on managerial training results.
- \* To review the literature on the application of the assessment centre method as a predictor of managerial training results.
- \* To conduct an empirical study to determine the validity of the MMAC as a predictor of training results achieved on the SMTc.

#### **1.3.5 The research question**

Based on the rationale for this study, and in support of the goals thereof, the research question is formulated as follows :

***IS THERE A RELATIONSHIP BETWEEN ASSESSMENT CENTRE PERFORMANCE AND MANAGERIAL TRAINING RESULTS?***

### **1.4 SCOPE OF THE REPORT**

#### **1.4.1 General orientation**

The present report will consist of two main sections. The first section, chapters 1-4, will consist of a literature survey regarding the prediction of training results, the application of the assessment centre method, and the use of assessment centres to predict training results. The second section, chapters 5-7, will report on an empirical study which was conducted to determine the validity of an existing assessment centre in predicting managerial training results.

#### 1.4.2 Specific chapters

Chapter 1 addresses the background to the study, its rationale and aim. Chapter 2 will entail a literature survey on the prediction of managerial training results. The content of current-day managerial training programmes, the use of training results as a criterion of training success, and the validity of instruments used as predictors of training results will receive attention. Chapter 3 will entail a literature survey of the assessment centre method. The assessment centre method will be defined, the underlying rationale on which it is based explained, and the predictive validity thereof investigated. Chapter 4 will investigate the use of assessment centres utilised specifically to predict managerial training results.

Chapter 5 will describe the methodology used for this study. Following the formulation of a hypothesis, a description of all measuring instruments will be given, which includes a discussion of the independent and dependent variables used in this study. The research design that was followed will then be discussed, followed by the data-analysis methods used. Chapter 6 will report the results obtained from the various statistical analyses conducted, while these results will be discussed in the final chapter of this study, namely chapter 7. The final chapter will also contain a discussion of the limitations of this study, conclusions reached, recommendations made and directions for possible future research.

#### 1.5 SUMMARY

In this chapter the background to the study was given which included a discussion of the importance of management for organisations as well as the necessity of training, developing and educating managers in

order to acquire the necessary knowledge and skills needed to cope with the requirements of the environment. The importance of selecting the right candidates to attend such managerial training was also emphasised.

The rationale for the study was shown to be the necessity of validating an assessment centre if it is utilised as a selection device such as determining entry to a training programme. Following this, the general aim of the study was formulated, being to determine whether the MMAC predicts training results achieved on the SMTA. Three specific objectives were also formulated.

The chapter was concluded by giving an overview of the scope of the report. It was shown that the report will consist of two sections and seven chapters, with section one covering the literature survey and section two describing the empirical study conducted to investigate the research problem.

In the next chapter, the prediction of managerial training results will be discussed. The use of managerial training results as a criterion of training success as well as the validity of various instruments commonly used as predictors of managerial training results will be discussed.

## CHAPTER 2

### THE PREDICTION OF MANAGERIAL TRAINING RESULTS

#### 2.1 INTRODUCTION

In this chapter the available research concerning the prediction of managerial training results will be surveyed. A brief overview of the subject-matter covered by current-day managerial training programmes will be given after which the use of managerial training results as a criterion of managerial training success will be discussed. The chapter will conclude with a discussion of the validity of various instruments commonly used as predictors of managerial training results.

#### 2.2 THE CONTENT OF CURRENT-DAY MANAGERIAL TRAINING PROGRAMMES

Managerial training and development programmes are implemented in most private and public organisations every year (Burke & Day, 1986; Saari et al, 1988; Wexley & Latham, 1981), and according to Goldstein (1980) and Wexley and Latham, the objective of most of these programmes is to teach or improve various managerial skills in order to improve on-the-job performance. In particular, a study by Burke and Day found that managerial training is focused on improving performance in the areas of general management, human relations/leadership, self-awareness, problem-solving/decision-making, and motivation/values.

#### 2.3 THE USE OF TRAINING RESULTS AS A CRITERION OF TRAINING SUCCESS

According to Sackett and Ryan (1985), training results are often used as criteria in predictive validity

studies due to convenience, as such measures are readily available. Gatewood and Feild (1990) claim that the use of training results as criteria is desirable, due to the control available during a training course in the form of consistently administered devices, while the legal permissibility of utilising training results as criteria is confirmed by The Guidelines (1992).

### 2.3.1 The measurement of training results

Cascio (1987) distinguishes between two types of training criteria used to measure training results, namely internal criteria (those which are linked directly to performance in the training situation), and external criteria (measures designed to assess actual changes in job behaviour). Kirkpatrick (1979) identified four levels in the evaluation of the effectiveness of training programmes: reaction to training, learning achieved, job behaviour, and organisational outcomes. According to Martin (1957), as quoted by Cascio (1987), reaction to training and learning achieved are internal criteria, and job behaviour and organisational outcomes are external criteria. Table 2.1 indicates the link between the levels of training evaluation, the type of training criteria, and the purpose for which each level and type are utilised.

Cascio (1987) further distinguishes between **qualitative criteria** (attitudinal and perceptual measures by administering, for example, written instruments), and **quantitative criteria** (for example, a written examination at the conclusion of training to assess mastery of the programme content).

**TABLE 2.1**  
*Levels of Training Evaluation, Type of Training Criteria, and Purpose of each type of Criteria*

Level of training evaluation	Type of criteria	Purpose
Reaction to learning	Internal	Performance in the training situation
Learning achieved		
Job behaviour	External	Actual changes in job behaviour
Organisational outcomes		

(Cascio, 1987; Kirkpatrick, 1979)

### **2.3.2 Learning achieved as a measure of training success**

Within the above framework the measurement of training results for use as criteria of training success would involve the qualitative and quantitative measurement of internal training criteria, such as a trainee's reaction to learning and learning achieved during or at the completion of a training programme. To determine if learning has taken place during a training programme, Kirkpatrick (1979) states that it is necessary to determine what principles, facts and skills a trainee has acquired during the training period. According to Kirkpatrick, the concern is not on-the-job application of these principles, facts and techniques, but the degree to which they were acquired during the training programme. Likewise, Camp et al (1986) argue that determining if learning objectives have been achieved is a means of measuring training success, as it provides evidence that a trainee has mastered the contents of a training programme.

### **2.3.3 Methods of quantifying learning achieved**

Gateway and Feild (1990) suggest three basic methods of quantifying all learning objectives, namely judg-

ments made by the training instructor about the trainee, scores on paper-and-pencil knowledge tests and performance test scores.

(a) Training instructor judgments

For Gatewood and Feild (1990), judgments made by the training instructor about the trainee, could be about parts of the training programme or a judgment of overall efficiency. Burke and Day (1986) claim that instructor ratings of management trainees during role play tests, play an important part in measuring various aspects of human relations proficiency. An example of a practical application of instructor judgment is reported by Tziner and Dolan (1982), who used instructor ratings of officer performance on practical command exercises as a training performance criterion.

The use of instructor judgments to measure training performance has not been without criticism. Gatewood and Feild (1990) point out that such data is viewed with scepticism because the evaluation is based on the opinion or judgment of the instructor, making it judgmental and susceptible to rater bias. In an attempt to ensure objectivity and reliability, various systematic approaches have been proposed. Cascio (1987) and Gatewood and Feild (1990), suggest rater training as an effective strategy while Banks and Roberson (1985) suggest the following two strategies: Build in as much structure as possible in order to minimise the amount of discretion exercised by a rater and do not require raters to make judgments that they are not competent to make. Kirkpatrick (1979) supports these suggestions, claiming that a well-planned and implemented instructor rating system can produce objective measures of the level of learning that has been achieved.

Despite the mentioned disadvantages, Cascio (1987), and Gatewood and Feild (1990), claim that the observation and judgment of training performance is unavoidable in modern business, as much of the learning that takes place, such as managerial and human relations skills, is no longer of the type that produces tangible, easily quantifiable performance. Rather than trying to limit this method, the above authors state that efforts should be increased to improve it.

(b) Scores on paper-and-pencil knowledge tests

Where principles and facts are taught rather than techniques, Kirkpatrick (1979) recommends the use of paper-and-pencil knowledge tests, by means of a systematic and statistically oriented approach, to obtain measures of the amount of learning that has taken place. This can be done by following what Kerlinger (1973) calls a Pretest-Posttest Design, where the same test is administered to all candidates prior to the training programme and again after the programme. According to Kirkpatrick (1979), a comparison of before and after scores can be made to prove statistically how much learning has taken place. This would then produce what Cascio (1987) calls a quantitative criterion, being a numerical measure of mastery of the training programme's content.

Examples of studies utilising objective paper-and-pencil knowledge tests to determine attainment of learning objectives are those reported by Borman (1982), in which three objective tests were used to measure mastery of prospecting and selling techniques, Feltham (1988), in which an academic grade of Police examination performance was used as a criterion, and Jones and Harrison (1982), in which a grade attained in an examination on theoretical subjects was used as

a criterion of training success. In reviewing studies utilising training performance as a criterion of success, Gaugler et al (1987), and Hunter and Hunter (1984), found the use of numerical grades obtained from paper-and-pencil knowledge tests to be the most prevalent in predictive validity studies.

(c) **Performance test scores**

According to Cascio (1987), a performance test usually represents a final evaluation of what a trainee can do at the completion of training, and focuses on present competence. Gatewood and Feild (1990) describe such tests as requiring a testee to complete some activity, either behaviourally, verbally or in writing, under structured testing conditions.

The primary objective of performance tests is to assess *the ability to do* and not to assess the ability to know. According to Robertson and Downs (1989), work sample performance tests assume the applicant already has the knowledge, ability, and skill required to perform the task, and this view is confirmed by Cascio and Phillips (1979) who state that the objective of a performance test is to assess developed skills. For this reason, Gatewood and Feild (1990) suggest that a performance test is particularly suited to determine the amount of learning that has taken place during a training programme, as it provides *direct* evidence of a trainee's knowledge, ability and skill to perform a specific activity. This is confirmed by numerous studies making use of performance tests to measure the attainment of learning objectives, such as underwater exercises for navy divers (Helmreich, Bakeman & Radloff, 1973; Ryman & Biersner, 1975), practical command exercises for officers (Borman, 1982; Tziner, 1984), practical military appreciations (Pagonis & Cruikshank, 1992), role playing exercises and business

games for managers (Burke & Day, 1986), and simulated interviews for personnel managers (Cuming, 1980). In all these studies, trainees were rated by instructors on a numerical scale and allocated a final training proficiency grade which was utilised as the criterion of training success.

#### **2.3.4 Summary**

According to the literature revised above, training results are widely used as a criterion of training success, as they are readily available, quantifiable, and, if utilised correctly, a reliable measure of the amount of learning that has taken place during a training programme. Training instructor judgments, scores on paper-and-pencil knowledge tests, and performance test scores can all be utilised to quantify learning objectives in order to determine the degree to which a trainee has mastered the content of a training programme.

### **2.4 TRAINING PREDICTORS AND THEIR VALIDITY**

According to Gatewood and Feild (1990), predicting future success is the essence of selection, and when a selection device is used to determine entry into a training programme, Robertson and Downs (1979) state that an attempt is made to predict the subsequent capacity of the applicant to master the contents of the training programme.

#### **2.4.1 Approaches to prediction**

For McCormick and Ilgen (1984), prediction in behaviour research typically requires the use and measurement of two types of variables, namely the predictor, and that which is being predicted, the criterion. When training success is predicted, the

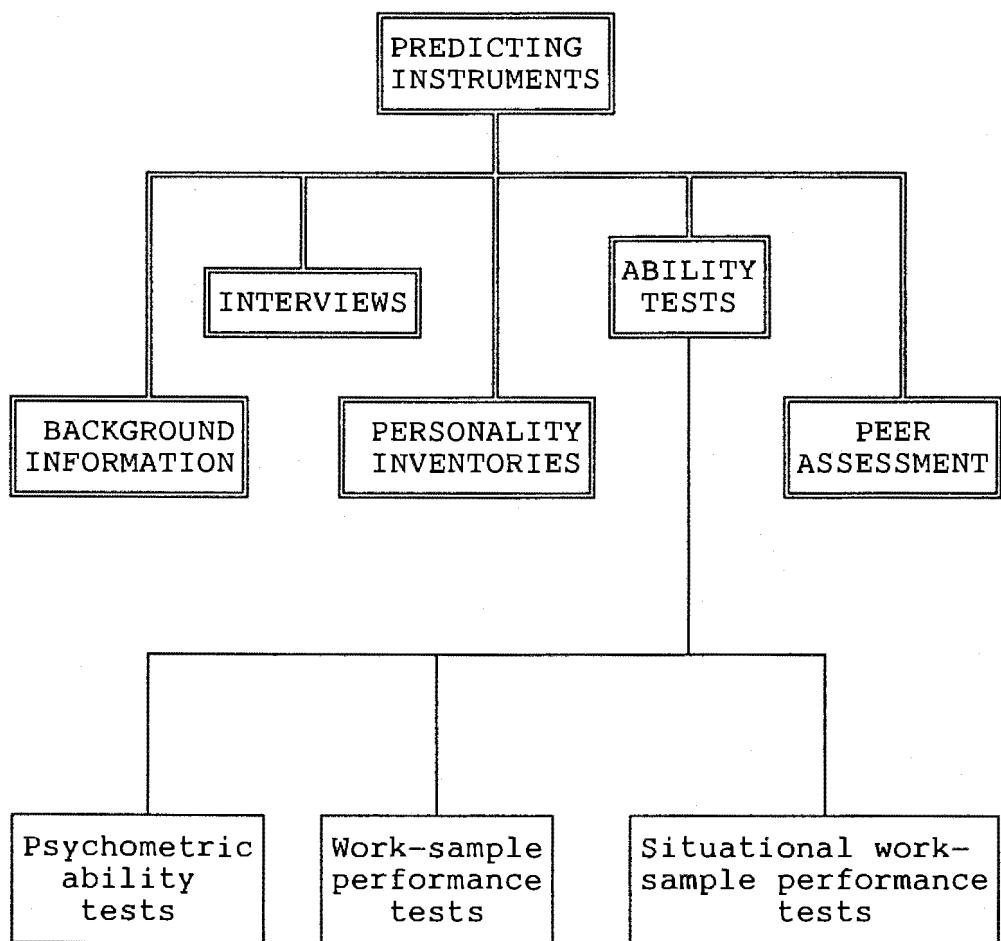
criterion will be some measure of training results, such as grades obtained at the completion of a training programme, while the predictor will be some or other measure used to select candidates to attend the training programme (Borman, 1982; Cascio, 1987; Robertson & Downs, 1979).

Wide use has been made of various traditional tests to predict training results (Ghiselli, 1973), and these tests have been referred to as signs of assumed underlying determinants of behaviour, which are used to indicate predispositions to behave in certain ways in the future (Cascio, 1987; Dreher & Sackett, 1981; Thornton & Byham, 1982; Wernimont & Campbell, 1968). An alternative to the traditional psychometric predictor or sign was offered by Wernimont and Campbell, who claim that conventional tests lack relevance. They argue that test items are often unrelated to the work required on the job for which an applicant is being considered, and that for effective selection, it would be more fruitful to focus on predictors that are as similar to the criterion as possible, such as samples of past or present behaviour. Asher and Sciarrino (1974) also argue the value of having what they describe as point-to-point correspondence between predictor and criterion space. This means that the more commonalities between predictor and criterion, the higher the validity is likely to be. For Taylor (1982), this represents a departure from the traditional assumption that predictor and criterion measures need to be different.

#### **2.4.2 Validity of instruments used to predict managerial training results**

A large volume of research exists in which a variety of instruments have been used to predict managerial

training results (Asher & Sciarrino, 1974; Beatty & Schneier, 1981; Cascio, 1987; Cascio & Phillips, 1979; Dunnette, 1976; Dreher & Sackett, 1981; Gatewood & Feild, 1990; McCormick & Ilgen, 1984; Thornton & Byham, 1982), and a schematic representation resulting from this research is given in Figure 2.1.



**FIGURE 2.1 : Instruments used to predict managerial training results**

**(a) Background information**

According to Cascio (1987), and Gateway and Feild (1990), application forms, life history questionnaires, reference checks and previous experience, are

all examples of background information. Cascio (1987) claims these instruments can be used to sample past or present behaviour, briefly but reliably, as the assumption is that aspects of an individual's total background are related to later success in a specific activity. This assumption is based on Wernimont and Campbell's (1968) behaviour consistency model, in which the best indicator of future performance is seen to be past performance. For these authors the implicit assumption of such questionnaires is that many background items represent an attempt to assess previous achievement which would predict future achievement. According to Asher and Sciarrino's (1974) point-to-point theory, such questionnaires can be successful if they contain information which has a high correspondence with the behaviour they are predicting.

In a study to test Wernimont and Campbell's (1968) consistency model, Jones and Harrison (1982) developed a Reference Report to predict the training performance of young men joining the Royal Navy as officers. The Reference Report contained scales emphasising personal qualities required by candidates to meet the demands of initial training. These Reports were sent to Headteachers of schools whose pupils apply for officer entry and completed for each applicant. The training criteria consisted of a theoretical examination mark, an "officer like qualities" mark, reflecting leadership and general conduct, and a total mark, being a combination of the previous two grades. A corrected correlation of 0,36 was obtained between the total of the seven Reference Report scales and the total mark achieved in training. Jones and Harrison (p. 40) describe this as being "around the level typically found for psychometric tests against training criteria", and "higher than would be expected on the basis of similar research". They conclude that

these results lend support to Wernimont and Campbell's (1968) consistency theory, as previous achievement successfully predicted future achievement.

In a meta-analysis of the validity of alternatives to ability tests, Hunter and Hunter (1984) reviewed the correlations obtained in studies using various background information items as predictors with training success as the criterion. The authors described training success as the grades obtained at the completion of the training programmes, but they did not indicate whether these were managerial or non-managerial programmes. The correlations recorded are reflected in Table 2.2, ranging from 0,23 for reference checks to 0,30 for biographical information and college grade. Hunter and Hunter conclude that although these validities are statistically significant, background measures should be used in combination with other predictors, such as ability tests, to ensure significant practical utility.

**TABLE 2.2**  
*Validities of Background Information Measures with Training Success as the Criterion*

Predictors	Correlation coefficients
Biodata	0,30
Reference checks	0,23
College grade	0,30

(Hunter and Hunter, 1984)

#### (b) Interviews

The interview is recognised as the most well-known and widely used selection technique, forming part of most

selection procedures (Gatewood & Feild, 1990; Glueck, 1982; Nunns & Kruger, 1986). According to Cascio (1987), the record of the employment interview as an aptitude test to predict future performance has been dismal, yet it has endured and flourished despite intense criticism.

Borman (1982) used a structured interview to measure achievement motivation, potential for being a self-starter, and commitment to the Army. He then used these measures as predictors of subsequent military recruiter training performance, which he considered to be managerial training performance. A correlation of 0,07 with a theoretical training examination score as the criterion, and 0,26 with a grade awarded for the practical demonstration of learned skills, was obtained.

Tziner and Dolan (1982) used a structured selection interview to evaluate a candidate's expected success in training as an officer, and obtained a correlation of 0,14 with performance in training, in the form of a final course grade as the criterion. In a 1 to 19 year follow-up of the validity of a police assessment centre used to assess managerial potential, Feltham (1988) obtained correlations of 0,09, 0,03 and -0,04 between three interviews as predictors and an overall training performance grade as the criterion.

When managerial training results are used as the criterion of interest, the above studies confirm Cascio's (1987) claim that the interview as a selection device has not fared well. A reason for this is offered by Borman (1982), who suggests that the interview does not effectively assess aspects which are important for success in officer training, such as cognitive ability and managerial behaviour.

The above research supports Wernimont and Campbell (1968) and Asher and Sciarrino (1974), who claim that higher validities will be obtained when the predictor and criterion are as similar as possible. Where the interview was used to predict theoretical training grades, Borman (1982) and Tziner and Dolan (1982) reported insignificant correlations of 0,07 and 0,14 respectively. When the interview was correlated with the practical demonstration of learned skills, Borman reported a significant correlation of 0,26. The significant correlation obtained with the practical criterion could be explained by the fact that the interview, being manifested verbal behaviour, is closer to a practical demonstration of learned skills than to a written theoretical test.

(c) Personality inventories

According to Gatewood and Feild (1990), the use of personality information in human selection is one of the most complex topics in the assessment of applicants. They argue that the reason for this is that many organisations use data about personality as an important part of the evaluation of applicants, while the documented usefulness of such information in making accurate selection decisions is limited.

Few published studies report on the ability of personality test scores to predict managerial training performance. However, in one such study, Borman (1982) investigated the validities of personality test scores and scores obtained on a personality/vocational interest battery as predictors of military recruiter training success, and found validity coefficients ranging from nonsignificant to very low. In another documented study, Tziner and Dolan (1982) conducted a study on female applicants who were applying for officer training, to determine which predictors of

performance in training had the highest validity. A predictor used, amongst others, was a personality inventory score, and a nonsignificant correlation of 0,06 was obtained. From these two reported studies, it can be tentatively concluded that personality inventories have little value as predictors of managerial training results.

(d) Peer assessment

Kane and Lawler (1978), in a comprehensive review of studies to clarify the strengths, weaknesses, and unique features of the various forms of peer assessment, distinguish between three assessment methods of ranking a person, as judged by the members of a well-defined group to which he belongs. These are peer nominations, peer rating and peer ranking. Only in the case of peer nominations was training performance used as a criterion, reported in Table 2.3.

Most of the studies reviewed by Kane and Lawler (1978) have been undertaken in military settings, and most of the characteristics assessed have to do with leadership ability. The authors consequently conclude that peer nomination is a valid predictor of training results, but that three aspects should not be overlooked, namely, that the validity of this method may largely be confined to leadership performance; that the method has a built-in bias favouring its validity; and that evidence for its validity does not establish its utility as a predictor compared to other predictors.

Reilly and Chao (1982) reviewed studies on the use of peer evaluation to predict training results, and found 10 coefficients that they could use to compute an average validity of 0,31 for all published studies.

TABLE 2.3  
*Validities of Peer Nominations with Training Performance as the Criterion*

Outcome assessed	Criterion	Validity
Qualifications to be the cadet group leader	Pass or fail flight training	0,27**
Same as above	Final average in preflight courses	0,50**
Officer potential	Final academic grade	0,42**
Interest in Navy	Final academic grade	0,23**
Likely to succeed as an officer	Final academic grade	0,72**

(Kane & Lawler, 1978)

\*\*p < 0,01

They explain this significant validity as being due to the fact that a peer has the opportunity to observe the behaviour of a trainee under a variety of conditions, which may be more revealing of future training performance. The writers do not, however, indicate if these results are for managerial or non-managerial training.

Tziner (1984) conducted a study on female applicants, accepted for officer training, to validate various predictors against performance in training, and found that the peer rating measure obtained prior to training was a highly valid predictor of the results obtained by trainees at the conclusion of the programme ( $r = 0,51$ ,  $p < 0,05$ ).

The available studies reviewed indicate that peer assessment is a valid predictor of managerial training results. For Reilly and Chao (1982), such validity can

be ensured if a thorough job analysis of the training programme is conducted. This would be to specify the important behavioural factors or dimensions required to succeed at training, and the assessment thereof would be based on clearly defined behavioural terms.

(e) **Psychometric ability tests**

A variety of paper-and-pencil ability tests have been used to predict future performance (Cascio, 1987; Gatewood & Feild, 1990; Hunter & Hunter, 1984; Korman, 1968), including tests that measure general intelligence; verbal, nonverbal, spatial relations and numerical ability; inductive reasoning; and perceptual speed and accuracy (Cascio, 1987).

Ghiselli (1973) conducted a comprehensive literature review regarding ability tests used to predict success in training. This was to determine trends in the validity of tests used for this purpose. He classified ability tests into four groups, namely intellectual abilities, spatial and mechanical abilities, perceptual accuracy, and motor abilities, and used success in training in managerial occupations as one of the main criteria. Ghiselli concluded that some validity coefficients were quite respectable for training criteria in managerial occupations, ranging from 0,02 to 0,30 as indicated in Table 2.4.

Siegel and Bergman (1975) used standard intellectual ability tests to predict training performance of United States Navy recruits. The recruits had to do various performance tests at the completion of training, and the scores obtained were used as measures of training success and served as the training criteria. The correlations obtained in the study are reflected in Table 2.5.

**TABLE 2.4**  
*Validities of Ability Tests used to Predict Training Results in Managerial Occupations*

Type of test	Criterion	Validity
Intellectual ability test	Training results	0,30
Spatial and mechanical ability test	Training results	0,28
Perceptual accuracy test	Training results	0,23
Motor abilities test	Training results	0,02

(Ghiselli, 1973). Note :  $p$  not available.

**TABLE 2.5**  
*Correlations between Mental Ability Tests and Training Performance*

Criterion (performance test)	Correlations
Practical task	0,17
Practical task	0,14
Knowledge test	0,47*
Alertness/common sense test	0,39*
Practical test (procedures)	0,33*

(Siegel & Bergman, 1975)

\* $p < 0,05$

In the discussion of their results, Siegel and Bergman (1975) note that the two practical tests, requiring mainly psychomotor skills, produced insignificant correlations of 0,17 and 0,14, while the tests requiring more intellectual skills produced significant correlations of 0,33, 0,39 and 0,47. According to these authors, the results support Asher and Sciarrino's (1974) point-to-point theory, as the

intellectual ability tests have a high degree of correspondence with the criterion measure requiring mainly intellectual ability to perform well, and as such, produced the highest validities.

Gordon and Kleiman (1976) used an intelligence test to predict training performance of Police recruits and obtained correlations of 0,33 and 0,56 with two training groups. An analysis of the study indicates that the elements measured by the intelligence tests are similar to the elements of the learning tasks on the training programme. According to Asher and Sciarrino's (1974) point-to-point theory, this similarity between predictor and criterion could explain the high validities obtained.

Mills (1976), in two studies conducted to determine the validity of selection methods for Cincinnati Police, used a general intelligence test as a predictor and performance in training at the Police Academy as the criterion. Correlations of 0,59 and 0,70 were obtained for the two studies. Mills explains these high predictive validities as being a result of the academic orientation of the Academy training, requiring mainly intellectual aptitude to successfully master the content of the course.

Tziner and Dolan (1982) conducted a study on female applicants, applying for officer training, in an attempt to determine which predictors of performance in training had the highest validity. Predictors used, amongst others, were verbal, non-verbal and inductive intelligence scores, which produced correlation coefficients of 0,39, 0,21 and 0,31 respectively. Tziner (1984) conducted a follow-up study on these same female applicants, who were accepted for officer training, to once again validate various predictors against performance in training,

and found that in contrast to the previous study, the verbal, non-verbal, and inductive intelligence scores used as predictors, now produced insignificant correlations of 0,12, 0,17 and 0,08 respectively.

A reason for these different validities reported in the above two studies could be due to the underlying essence of the performance criterion used to validate the intelligence scores (Cascio, 1987). In the Tziner and Dolan (1982) study, scholastic aptitudes as measured by the intelligence tests were important for performance in the officer training, as the training consisted of various aspects which placed intellectual demands on the trainees. In the Tziner (1984) study, however, training consisted of, amongst others, various command exercises, and, as Tziner points out, command abilities are presumably much more crucial to success in such officer training than the pure scholastic aptitudes measured by the intelligence tests, and hence the low predictive validities obtained for these scores. The results obtained in the Tziner and Dolan (1982) and Tziner (1984) studies support Asher and Sciarrino's (1974) point-to-point theory, which states that the highest validities will be obtained when there is a high degree of correspondence between predictor and criterion measures.

**(f) Work-sample performance tests**

According to Glueck (1982), a work-sample performance test is an experience which involves actually doing a sample of the work involved in the job in a controlled situation, and Cascio and Phillips (1987) state that the primary objective of such a test is to assess the ability *to do* rather than the ability *to know*.

Work-sample performance tests are based on what Wernimont and Campbell (1968) refer to as a beha-

vioural consistency approach to selection. In this approach they suggest that the predictor used should be as similar as possible to the criterion of interest. For this reason, these tests are designed to approximate job content closely, requiring applicants to perform actual segments of the job in a controlled situation (Cascio, 1987; Cascio & Phillips, 1987; Glueck, 1982; Gatewood & Feild, 1990; Nunns & Kruger, 1986). Asher and Sciarrino (1974) argue that work-sample performance tests are more consistent with behaviour required on the actual job, and as such, there is a point-to-point correspondence between predictor (work-sample test) and criterion (future behaviour), which will ensure high validities.

Although work-sample performance tests have been used extensively as selection devices (Cascio & Phillips, 1987; Gatewood & Feild, 1990; Hunter & Hunter, 1984; Robertson & Kandola, 1982), relatively few studies are available where managerial training results have been used as a criterion when validating such tests. Studies that have utilised managerial training criteria include those by Asher and Sciarrino (1974), Gordon and Kleiman (1976), and Robertson and Kandola (1982).

In the first reported study, Asher and Sciarrino (1974) selected a variety of work-sample performance tests and classified them into one of two categories. Motor work-sample tests were considered to be those in which subjects had to physically manipulate things, while verbal work-sample tests were considered to be those requiring individuals to cope with people-oriented or language-oriented problems. One of the criteria used in this study was success in training, and the specific measures used were either the grades obtained at the completion of training, the grades achieved during training, or a scale rating of train-

ing performance. Asher and Sciarrino (1974) obtained a correlation coefficient of 0,55 for the verbal work-sample tests and 0,45 for the motor work-sample tests, and conclude that work-sample tests are effective predictors of success in certain types of training.

In another study utilising managerial training results as criteria, Gordon and Kleiman (1976) conducted an investigation to provide a direct comparison of the validities of an aptitude (intelligence) test and a work-sample test, in forecasting training performance. The criterion of training success consisted of the sum of scores obtained on all graded exercises at a Police training academy. Three groups of subjects in the academy were given the work-sample test and the intelligence test, and the correlations obtained are reflected in Table 2.6.

TABLE 2.6  
*Correlations of Work-sample and Intelligence Tests with Training Performance as the Criterion*

Predictor	Correlations		
	Group 1	Group 2	Group 3
Work-sample test	0,52**	0,72**	0,64**
Intelligence test	0,33	0,15	0,56**

(Gordon & Kleiman, 1976)

\*\*p < 0,01

As a result of the study, Gordon and Kleiman (1976) conclude that work-sample tests are better predictors of Police academy training results than an intelligence test. The authors claim that the high validities obtained support Asher and Sciarrino's (1974) point-to-point theory, as the work-sample test had a high correspondence with the training criterion.

In a follow-up review on validity data and work-sample performance tests, Robertson and Kandola (1982) classify work-sample tests into four categories, namely psychomotor tests, individual decision-making tests, job-related information tests, and group decision-making tests. Using data from over 60 different validation studies and training criteria such as examination performance and instructor ratings at the end of training, the authors obtained the correlations as reflected in Table 2.7.

**TABLE 2.7**  
*Validity Coefficients for Different Types of Work-Sample Tests with Training Performance as the Criterion of Success*

Type of test	Median of correlations
Psychomotor	0,38
Individual decision-making	0,25
Job-related information	0,50
Group decision-making	Not available

(Robertson & Kandola, 1982) Note : p not available.

Robertson and Kandola (1982) conclude that the results suggest work-sample tests are valid predictors of training performance, but that available data does not allow for generalisations about the value of work-samples compared with conventional predictors such as certain paper-and-pencil tests.

From the above research the use of work-sample performance tests as predictors of managerial training results has been shown to produce significant validities. For Gatewood and Feild (1990), these high validities are possible if care is taken in the construction of these tests to ensure that they are representative of job activities.

## (g) Situational work-sample performance tests

Situational work-sample performance tests differ from normal work-sampling, in that they use *simulations* of job content, rather than a sample of *actual* job content (Beatty & Schneier, 1981; Cascio & Phillips, 1987; Howard, 1974; Thornton & Byham, 1982). For Robertson, Gratton and Sharpley (1987), there is also a marked difference in the design procedure for the two tests when used for managerial selection purposes. While normal work-sample tests are designed directly from a job analysis to assess managerial performance behaviour, situational work-sample tests are developed to elicit abstract managerial behavioural dimensions which are first identified from a job analysis.

Moses (1977) says that situational tests simulate an actual situation which requires critical behaviour related to success on the job, so that each participant is given the same opportunity to demonstrate his abilities in standardised situations. According to Jaffee (1971), situational tests attempt to duplicate the critical aspects of a job and create simulations that resemble those in which a manager frequently finds himself. Boche (1977) and Robertson et al (1987), say that proper situational tests are work-samples of the managerial role, and as such can be used to predict performance in the actual managerial role.

In propagating the use of situational tests as selection devices, Weislogel (1954) claims that these tests have specific advantages over other test techniques, as they permit evaluation of three interrelated aspects of typical performance:

- Possession of the necessary skills;

- Recognition of the need to apply these skills at the appropriate time; and
- Motivation, or willingness to apply the skills.

As such, Weislogel (1954) claims that the situational test lends itself well to settings in which the job performance involves a variety of activities, including interactions with other individuals and actions based on the exercise of judgment.

In the area of managerial selection, two types of situational tests are commonly used (Cascio, 1987; Greenwood & McNamara, 1967; Gatewood & Field, 1990). In *group exercises*, participants are placed in a situation where the successful completion of a task requires interaction among the participants. In *individual exercises*, participants complete a task independently.

Various researchers (Mills, 1976; Tziner, 1984; Tziner & Dolan, 1982) investigated the validity of group and individual situational work-sample tests as predictors of training results. In a study to determine the validity of measures used to select Police officers, Mills (1976) compared test scores on two situational tests with performance in the Police Academy (training results) for two groups of recruits, and obtained the correlations reflected in **Table 2.8**.

For Mills (1976), the correlations obtained with the Clues situational test are highly predictive of performance at the Police academy. He ascribes these significant validities, and the insignificant validities obtained with the Foot Patrol Test, being due to the nature of the academically oriented Police academy training criterion. Mills explains that it requires cognitive ability to perform well on the

**TABLE 2.8**  
*Validity Coefficients of Situational Tests with Police Academy Training Results as Criteria*

Type of situational test	1964 Group (n = 42)	1965 Group (n = 15)
Foot Patrol Test	0,14	0,16
Clues Test	0,38*	0,43**

(Mills, 1976)

\*  $p < 0,05$

\*\*  $p < 0,01$

Clues Test, while the Foot Patrol Test requires applied behavioural skills. As such there was what Asher and Sciarrino (1974) refer to as a greater point-to-point correspondence between the Clues Test as predictor and the cognitive training criterion, than with the practical Foot Patrol Test as predictor and the cognitive training criterion, accounting for the higher validities obtained in the former case.

Tziner and Dolan (1982) conducted a study to determine the validity of various assessment centre measures used to predict the training results of female officers in the military, and report significant validities for various situational performance tests as reflected in Table 2.9.

These authors describe the results as indicating that the situational tests satisfactorily relate to performance in training. They argue that the significant validities are due to the fact that the situational exercises are tailor-made to allow for measurement of the appropriate dimensions, identified with a rigorous job analysis of the female officer role. Tziner and Dolan (1982) emphasise that the situational exercises were specifically designed to predict success in training, and not on-the-job performance.

**TABLE 2.9**  
*Correlations between Situational Tests and Performance  
in Officer Training*

Type of situational test	Validity coefficients
In-basket	0,18*
Oral presentation	0,19*
Field command game	0,25**
Role playing	0,34**
Leaderless group discussion	0,24**

(Tziner & Dolan, 1982)

\* p < 0,05

\*\*p < 0,01

Borman (1982) conducted a study to determine the validity of various measures used to predict the training results of military recruiters, whom he considered to be managers. These results are reported in Table 2.10, and from his study, Borman concludes that situational tests are reasonably successful in predicting recruiter training performance.

**TABLE 2.10**  
*Correlations between Situational Tests and Performance  
in Military Recruiter Training*

Type of situational test	Validity coefficients
Cold calls	0,33*
Interviews	0,34*
Interview with parent	0,41**
In-basket	0,33*
Speech	0,32*

(Borman, 1982)

\* p < 0,05

\*\*p < 0,01

Tziner (1984) conducted a study in which a composite situational exercise score was compared to the training performance of military female officers, and obtained a correlation of 0,76. The situational exercise score was derived by averaging the global evaluations attained on each of the situational tests, and although Tziner does not specify which tests were involved, they are the same as those reported by Tziner and Dolan (1982), as this study was a follow-up to the Tziner and Dolan study (See Table 2.9). Tziner concluded that "situational tests prove the most realistic testing ground for the demands of successful performance in managerial occupations - in this case, as an officer" (p. 158). He explains the high validities as being due to the fact that the situational exercises are purpose-made to measure the assessment dimensions necessary to succeed in Female officer training.

The studies by Borman (1982), Mills (1976), Tziner (1984), and Tziner and Dolan (1982) show that situational performance tests, as predictors of managerial training results, produce significant validities. The highest validities obtained are those in which the researchers claim that *the situational tests were specifically designed to predict training results*. This confirms the claim by Wernimont and Campbell (1968) and Asher and Sciarrino (1974), who state that the highest correlations will be obtained if the predictor (situational test) is as similar as possible to the criterion of interest (training results), that is, if the behaviour required on the situational test is consistent with the behaviour required during the training course.

## 2.5 SITUATIONAL TESTS AND THE ASSESSMENT CENTRE METHOD

The use of simulation or situational work-sample exercises is one of the main distinguishing features of the assessment centre method where they are used to assess the presence of various dimensions of behaviour in candidates (Cascio, 1987; Gatewood & Feild, 1990; Thornton & Byham, 1982). Although Thornton and Byham found that many centres also use paper-and-pencil tests with situational exercises as part of their overall assessment process, which are then combined to arrive at an overall assessment rating (OAR) for each candidate, Crooks (1977) states that situational exercises are the main instruments used in the assessment centre method to determine the OAR (See Table 3.1). Various meta-analyses of assessment centre research (Gaugler et al, 1987; Schmitt et al, 1984) show that the OAR, which, according to Blanksby and Iles (1990) is the predictor most used in assessment centre validity studies, is a valid predictor of managerial performance.

From the research regarding the validity of situational tests in predicting training results (Borman, 1982; Mills, 1976; Tziner, 1984; Tziner & Dolan, 1982), the fact that situational exercises are one of the main distinguishing features of the assessment centre method (Cascio, 1987; Gatewood & Feild, 1990; Thornton & Byham, 1982), the extensive use of situational exercises to determine the OAR (Crooks, 1977), and the validity of the OAR in predicting managerial performance (Gaugler et al, 1987; Schmitt et al, 1984), it could be possible that *the assessment centre method is a valid predictor of managerial training results*. To test this possibility, the assessment centre method as a predictor of overall managerial performance will be investigated in chapter 3, after which the use of the assessment centre to

predict managerial training performance in particular will receive attention in chapter 4.

## 2.6 SUMMARY

In this chapter the prediction of managerial training results was discussed. The content of current-day managerial training programmes was reviewed and the use of training results as a criterion of training performance was investigated. The validity of various training predictors was discussed, and from this discussion it emerged that reference reports, cognitive ability tests, peer assessment, verbal work-sample performance tests and situational work-sample performance tests show the most potential of being used as predictors of managerial training results. Interviews, personality inventories, and motor work-sample tests produced either low or insignificant validities as predictors of managerial training results. The chapter was finally concluded by a discussion of the role of situational work-sample performance tests in determining an overall assessment rating (OAR) in an assessment centre. From this it emerged that the assessment centre method could be a valid predictor of managerial training results.

In the next chapter the assessment centre method as a predictor of overall managerial performance in various organisational settings will be investigated.

CHAPTER 3

## THE ASSESSMENT CENTRE METHOD

### 3.1 INTRODUCTION

In this chapter the assessment centre method will be discussed. The method will be defined, the underlying rationale on which it is based explained, the instruments and exercises most commonly used described, and the industrial applications thereof discussed.

The assessment centre as a predictor of overall managerial performance will also receive attention. The criteria utilised in assessment centre validity studies as well as the measures used as predictors will be discussed. The chapter will finally be concluded with an investigation of various assessment centre predictive validity studies, both overseas and in South Africa.

### 3 . 2 DESCRIPTION OF THE ASSESSMENT CENTRE METHOD

### **3.2.1 Assessment centres defined**

An assessment centre is a procedure which uses multiple assessment techniques to evaluate employees for a variety of manpower purposes and decisions (Thornton & Byham, 1982). Multiple trained observers are used and judgments about behaviour are made, in part, from specially developed assessment simulations. These judgments are pooled by the assessors at an evaluation meeting, during which all relevant assessment data is reported and discussed, and a consensus overall evaluation as well as consensus evaluations of each dimension are made (Task Force on Assessment Centre Guidelines, 1989).

### 3.2.2 The assessment centre approach versus psychometric prediction

The assessment centre method is based on what Wernimont and Campbell (1968) call samples of behaviour. Samples of behaviour are complex measures which require complex behaviour responses very much like the responses required on the job. Various simulation exercises are presented to candidates and the focus is on the relevant behaviours displayed by them, while the traits that presumably underly these behaviours are of less or no importance (Thornton & Byham, 1982). VJP

*what type →  
of responses*

Psychometric prediction is based on signs of behaviour, which are tests, questionnaires, or other paper-and-pencil instruments which measure predispositions to behave in certain ways. The actual behaviour involved in a sign may be nothing more than a minor physical response of marking a piece of paper (Thornton & Byham, 1982).

According to Thornton and Byham (1982), the process of making predictions of future behaviour is also quite different in assessment centres than in the psychometric approach. The latter combines the data from tests and predicts behaviour on the basis of the best statistical relationship among the variables. In assessment centres, the overall prediction is made after subjective consideration of the assessment data and judgmental combination of the variables.

### 3.2.3 Underlying rationale of assessment centres

The main underlying rationale of assessment centres is a behavioural or psychosituational approach which involves the direct measurement of

an individual's responses to various life-like or work-related situations (Berg et al, 1980). It is thus accepted that a situation can be "standardised" to assess individuals, so that more control of ratings can be gained by eliminating possible effects on the ratings caused by raters and different situational variables. An individual's responses to various life-like or work-related situations which can be created, can thus be measured in a more objective manner (Gatewood & Feild, 1990; Howard, 1974).

According to Thornton and Byham (1982), the assessment centre approach also assumes that people possess relatively enduring characteristics which influence their behaviour in various settings. By simulating certain job requirements, an opportunity is provided to evaluate skills and abilities which would otherwise not be available from observation of performance on the current job. By emphasising behavioural observation in work-sample and job simulation measurement techniques, it is possible to assess skills and abilities which are important for a new job or position before an individual actually fills it.

### **3.2.4 Instruments and exercises used in assessment centres**

*approach*

Thornton and Byham (1982) claim that managerial assessment centres have gone beyond the task and trait approaches to identify behavioural dimensions of managerial work. In this sense, Thornton and Byham describe a dimension as a cluster of behaviours which are specific, observable and verifiable. According to Cascio (1987), various assessment techniques are designed to elicit a sample of behaviours relevant to the

identified behavioural dimensions in the target job, while assessors then rate performance of the candidate on the dimensions. Byham and Thornton (1970) state that it is this use of multiple assessment techniques which is one of the strengths of assessment centres.

Crooks (1977) gives a list of typical instruments used in assessment centres which include paper-and-pencil tests, in-basket exercises, management games, leaderless group discussions, interviews, analysis exercises, and writing exercises. This list is expanded by Thornton and Byham (1982) who reviewed approximately 500 centres and reported the use of various assessment centre techniques in operating managerial assessment centres. These techniques are displayed in Table 3.1

### 3.2.5 Industrial applications of assessment centres

Since the inception of the first operational assessment programme in 1958 by AT&T to select first-level foremen (Moses, 1977), the method has been used over the years for various purposes.

*uses*

These include predicting success in management (1) jobs (Byham, 1970; Howard, 1974), predicting promotions (Klimoski & Strickland, 1977), the selection of salesmen (Bray and Campbell, 1968), predicting women's advancement into middle management (Ritchie and Moses, 1983), predicting training performance (Borman, 1982), and predicting Black leadership potential in South Africa (Charoux, 1990).

**TABLE 3.1**  
*Frequency of use of Assessment Centre Instruments*

Assessment technique	Percentage of frequency of use
In-Basket	95
Assigned-role leaderless group discussion	85
Interview simulation	75
Non-assigned-role leaderless group discussion	45
Scheduling	40
Analysis	35
Management games	10
Background interview	5
<u>Paper-and-pencil tests</u>	
- Intellectual	2
- Reading	1
- Mathematics and arithmetic	1
- Personality	1
Projective tests	1

(Thornton & Byham, 1982)

### 3.3 THE ASSESSMENT CENTRE AS A PREDICTOR OF MANAGERIAL PERFORMANCE

The increasing popularity of the assessment centre method has stimulated a large amount of research regarding its effectiveness as a predictor of managerial performance (Gaugler et al, 1987; Schmitt et al, 1984). In fact Moses (1987) argues that one of the major reasons why the method has survived and flourished, is that it rests on a strong research base. Regarding assessment centre research, Sacket and Ryan (1985) argue that the value of a particular centre in making predictions

is likely to depend on the criterion measures used to validate it.

### 3.3.1 Criteria utilised in assessment centre validity studies

In a meta-analysis of assessment centre validity, Gaugler et al (1987) found five main categories of criteria most commonly used in assessment centre validity studies, with individual criteria under each category. These categories are summarised in Table 3.2.

**TABLE 3.2**  
*Categories of Criteria Utilised in Assessment Centre Validity Studies*

Category	Criteria within category	Number of studies	Correlation coefficients
Job performance ratings	An overall performance rating. Field observation of performance.	28	0,36
Measures of career advancement	Change in salary. Number of promotions.	25	0,36
Ratings of potential	A rating of a manager's potential.	9	0,53
Measures of performance in training	Scores obtained in tests.	7	0,35
Dimension ratings	Rating a manager's performance on the dimensions used in the assessment centre.	5	0,33

(Gaugler et al, 1987)

The results reported in Table 3.2 show that:

- \* Most studies utilised either job performance ratings or measures of career advancement as criteria.
- \* The average validity coefficients for the prediction of different criteria range from 0,33 to 0,53, supporting Sackett and Ryan's (1985) claim that assessment centre validity depends on the criterion measures used.

### 3.3.2 Assessment centre measures used as predictors

In a study by Blanksby and Iles (1990), it was found that most of the research on assessment centres has concentrated on determining the validity of the Overall Assessment Rating (OAR) as the main predictor of various aspects of managerial performance. According to Thornton and Byham (1982), the OAR may be expressed in terms of level of potential, likelihood of advancement, or quality of performance in assessment. The OAR has also been expressed as a final judgment as to a candidate's acceptability for a specific post, for example, more than acceptable to unacceptable (Bray & Campbell, 1968), as the likelihood of advancing to a certain managerial level, for example, yes or no (Bray & Grant, 1966), or as an overall potential of success as a manager, for example, on a 5-point scale (Schmitt et al, 1984).

Various studies have been conducted where other assessment centre ratings, besides the OAR, have been utilised as predictors of managerial performance. Borman (1982) used the ratings obtained in individual assessment centre exercises as well as the OAR as a predictor of managerial performance, and Tziner and Dolan (1982) used ratings obtained in exercises, the OAR, as well as

an average of dimension ratings, as predictors of managerial performance.

### 3.3.3 Predictive validity research results

A large number of studies on assessment centre validity have been published in the literature, and meta-analyses of these research results have been conducted by Gaugler et al (1987), Hunter and Hunter (1984), and Schmitt et al (1984). Their reported mean predictive validities for assessment centres are 0,37, 0,43, and 0,41 respectively. Commenting on these results, Klimoski and Brickner (1987) conclude that assessment centres do indeed work and are useful tools for predicting the future success of potential managers.

Assessment centres have been utilised successfully for various other purposes besides the prediction of managerial potential. Bray and Campbell (1968), utilised the method to select salesmen. They evaluated newly hired candidates for sales positions by means of an assessment centre consisting of paper-and-pencil tests and assessment centre simulations exercises. These predictors were compared some months later with field performance ratings. The OAR produced a correlation of 0,51, while the four separate paper-and-pencil tests produced correlations of 0,25, 0,26, 0,28, and 0,02 respectively. These findings support the contention that assessment centres can be used for various personnel selection purposes.

As previously mentioned, the assessment centre method has also been utilised to predict women's advancement into middle management (Ritchie & Moses, 1983). In 1973 and 1974, over 1600 women

were assessed in a management assessment centre during which a prediction was made regarding their middle-management potential. After an average of 7 years, it was determined whether or not these women had reached middle management, and a correlation coefficient of 0,42 was obtained between the assessment centre prediction and subsequent managerial progress.

The assessment centre also appears to have long-term predictive validity (Bray & Grant, 1966; Hinrichs, 1978; McEvoy & Beatty, 1989; Mitchel, 1975). In a longitudinal study by Mitchel, data on 254 managers attending an assessment centre were examined for changes in validity over time. Assessment centre ratings were correlated with a criterion of salary growth, measured 1, 3 and 5 years after the managers were assessed. The results show that the validity coefficient did in fact increase over time.

Various studies have also been conducted to compare the validity of assessment centres to other predictors. Borman (1982), compared the validity of assessment centre ratings with the validity of other techniques. He compared the following as predictors of managerial training performance : first impression ratings, physical attractiveness and likeability ratings, a structured interview, a personality inventory, ratings of performance in individual assessment centre exercises, an overall consensus rating, and another overall rating obtained not by consensus but by simply adding up the ratings on all dimensions across all exercises. While a composite of assessment ratings yielded a corrected correlation coefficient of nearly 0,50, all the other predictors correlated near zero with training

performance, confirming the superiority of the assessment centre over other predictors when the criterion is training performance.

In South Africa a number of studies have been undertaken on various aspects of the assessment centre method (Augustyn & Van Wyk, 1988; Spangenberg et al, 1989; Sakinofsky & Raubenheimer, 1982; Stroebel & Raubenheimer, 1983). The Spangenberg et al study was conducted to determine the validity of a middle management assessment centre. The predictors used were five biographical items (grade, age, language, sex and race), and 13 assessment centre dimensions, while 19 Behavioural Anchored Rating Scales (BARS) and a Total Rating were used as criteria. A multiple correlation coefficient of 0,37 between assessment centre dimensions and the Total Rating was recorded, while nine of the 13 assessment centre dimensions correlated significantly with either the Total Rating or the majority of the BARS. The importance of this study is that it provided solid evidence of the ability of assessment centres to predict on-the-job performance, contradicting recurring criticism that assessment centres are more predictive of advancement criteria than performance criteria (Klimoski & Strickland, 1977; Turnage & Muchinsky, 1984; Klimoski & Brickner, 1987).

### **3.3.4 Discussion**

The large amount of available research provides conclusive evidence that assessment centres are valid predictors of various criteria of interest, including managerial success, advancement, job performance, training results, and managerial potential. Research also shows that assessment

centres are better predictors than various paper-and-pencil tests, personality inventories, and interviews. According to Sackett and Ryan (1985), final conclusions about the value of a particular assessment centre are likely to depend first on the criterion measures used, and secondly on the soundness of the methodology followed in the validity study. However, due to differences in organisations and circumstances, Thornton and Byham (1982) argue that it would be best to validate each assessment centre by concentrating on the specific aim for which it was established.

### **3.4 SUMMARY**

In this chapter, the assessment centre method was discussed. The method was defined, its underlying rationale explained, the instruments and exercises most commonly used described, and the industrial applications thereof discussed.

The assessment centre as a predictor of overall managerial performance also received attention. The criteria utilised in assessment centre validity studies as well as the measures used as predictors were discussed. The chapter was finally concluded by reviewing various assessment centre predictive validity studies, both overseas and in South Africa, which provided conclusive evidence that assessment centres are valid predictors of various criteria of interest, including training results.

In the next chapter, the use of assessment centres to predict managerial training results will be investigated. This will entail a discussion of

various studies in which managerial training results were utilised as the criterion measure to determine the predictive validity of the assessment centre.

## CHAPTER 4

### THE ASSESSMENT CENTRE AS A PREDICTOR OF MANAGERIAL TRAINING RESULTS

#### 4.1 INTRODUCTION

In this chapter the use of the assessment centre as a predictor of managerial training results will be discussed. The available research in which managerial training results were used to validate assessment centres will be surveyed, and the possible reasons for the validity results obtained will also be investigated.

#### 4.2 STUDIES IN WHICH MANAGERIAL TRAINING RESULTS WERE UTILISED TO VALIDATE ASSESSMENT CENTRES

In a review of available assessment centre research, seven studies were found where managerial training results were used to validate a variety of assessment centres, as summarised in Table 4.1. Although the target group is referred to as "Officers" and "Administrators", a study of the literature on the role of managers, leaders, officers and administrators indicate that they all manage as part of their functions (Bateman & Zeithaml, 1993; Borman, 1982; Cloete, 1986; Drucker, 1968; Hall & Donnell, 1990; Hanekom, 1987; Hersey & Blanchard, 1993; Kroon, 1990; Pagonis & Cruikshank, 1992; Peters & Waterman, 1982). As such, when the term "Officer" or "Administrator" is used, the term manager is implied, and when the term "Officer training" is used, managerial training is implied.

**TABLE 4.1**  
*Assessment Centre Validity Studies using Training Results as Criteria*

Author(s)	Purpose of assessment centre	Target group
Borman (1982)	Selection	Army officers
Feltham (1988)	Selection	Police officers
Gardner & Williams (1973)	Selection	Navy officers
Mays (1972)	Selection	Police officers
Tziner (1984)	Selection	Army officers
Tziner & Dolan (1982)	Selection	Army officers
Vernon (1950)	Selection	Civil service administrators

#### 4.2.1 Study 1

In a short-term follow-up study to determine the validity of Civil Service Selection Board procedures used to select Civil Service Administrators in the UK (Vernon, 1950), a correlation of 0,42 was obtained between the Final Selection Board Mark of the Extended Interview (the Civil Service name for an assessment centre), and the grades which trainees obtained on their initial training course. This correlation was significantly higher than the four correlations obtained between assessment centre performance and subsequent job performance, and Vernon (1950) concludes that this specific assessment centre was a valid predictor of officer training results, but a poor predictor of job performance.

#### **4.2.2 Study 2**

In a 25-year follow-up of an Extended Interview selection procedure used to select Royal Navy Officers, Gardner and Williams (1973) found that Interview Board Marks (that is, assessment centre exercises, paper-and-pencil tests, and biographical information combined), correlated significantly only with Engineer officer training results ( $r = 0,21$ ). However, when the results of the Civil Service Commission Examinations (CSCE) were combined with the Interview Board Marks and used as a composite predictor, correlations of 0,37 and 0,39 were obtained for Seaman and Engineer training results respectively. Thus, on its own, the assessment centre was not a good predictor of officer training results, but when combined with other predictors, its validity rose sharply. Another aspect of interest is that various cognitive ability tests produced correlations of 0,31 and 0,46 with Engineer and Seaman training results as the criterion respectively. These results appear to indicate that when the content of a managerial training course is academically oriented, cognitive ability tests can be utilised as valid predictors of training results.

#### **4.2.3 Study 3**

In a study to determine the validity of an assessment centre designed to help select United States Army recruiters, Borman (1982) used six methods to predict success in officer training, namely, first impression, physical attractiveness, and likeability ratings; ratings based on a structured interview; test scores on a personality

inventory and on a personality/vocational interest battery; ratings of subjects' performance in individual exercises; a statistical composite of assessment ratings (dimensions); and consensus ratings of subjects' performance after a day's activities. The criterion, success in training, consisted of a Phase 1 composite score (objective knowledge tests), and a Phase 2 score (application of skills learned). The correlations obtained are presented in **Table 4.2**.

For Borman (1982), these results indicate that the assessment centre approach is reasonably successful in predicting officer recruiter training performance. After correction for restriction of range, composite assessment ratings produced a correlation of near 0,50 with success in training. Ratings on the individual situational exercises also show reasonably high validities, ranging from a correlation of 0,32 to 0,41. What is also clear is that all the assessment centre ratings are superior predictors of managerial training results when compared to various traditional predictors which are usually utilised as selection devices. According to Borman, these high validities are due to the fact that training results and not on-the-job performance were used as the criterion of interest.

#### **4.2.4 Study 4**

Tziner and Dolan (1982) conducted a study to compare traditional selection predictors with various assessment centre predictors of success. The criterion used was performance in training, which was the final grade obtained by candidates upon completion of officer training. The specific

predictors used as well as the correlations obtained are given in **Table 4.3.**

**TABLE 4.2**  
*Validity of Various Assessment Centre Predictors with Training Performance as the Criterion*

Predictors	Training criteria	
	Phase 1 score	Phase 2 score
<u>Measures</u>		
First impression	0,23	-0,18
Physical attractiveness	-0,06	-0,20
Likeability	0,06	0,04
Structural interview	0,07	0,26
<u>Performance in individual exercises</u>		
Cold calls	0,33*	-0,03
Interviews	0,34*	0,15
Interview with parent	0,41**	0,24
In-basket	0,33*	0,36*
Speech	0,32*	0,27
Consensus rating	0,38**	0,27
<u>Personality/Vocational Interest Battery (Overall measure)</u>		
Personality Inventory	0,04	0,05
Ratings on dimensions (Figures in brackets: after correction for restriction of range)	(0,53)**	(0,50)**

(Borman, 1982)

\* p < 0,05

\*\*p < 0,01

**TABLE 4.3**  
*Validity of Traditional Selection Predictors  
 Compared to Assessment Centre Predictors with  
 Training Results as Criteria*

Predictors	Validities
Superior evaluation	0,16*
Selection interview	0,14*
Personality evaluation	0,06
<u>Intelligence tests</u>	
Verbal intelligence score	0,39**
Nonverbal intelligence score	0,21**
Inductive intelligence score	0,31**
<u>Individual exercises</u>	
In-basket	0,18*
Oral presentation	0,19*
Field command game	0,25**
Role playing	0,34**
Leaderless group discussion	0,24**
Consensus rating on dimensions	0,36**
OAR	0,38**

(Tziner & Dolan, 1982)

\* p < 0,05

\*\*p < 0,01

Tziner and Dolan (1982) claim that the results of this study suggest that the assessment centre is a good predictor of officer training results. The results also show that all aspects of the assessment centre, namely, individual exercises, dimension ratings and OAR, are superior to other predictors of training results, except for the intelligence scores which surpass all assessment

centre ratings with a correlation coefficient of 0,39. Tziner and Dolan (1982) state that a reason for this can be found in the fact that the criterion was training results, which in this case had a high academic content, and, as Hunter and Hunter (1984, p. 80) state, "*There is no job for which cognitive ability does not predict training success.*" Another reason can be found in the fact that the predictor (intelligence) was directly related to the criterion (graded academic training results), ensuring a point-to-point correspondence which, according to Asher and Sciarrino (1974), should ensure higher validities than when the predictor and criterion are unrelated.

#### 4.2.5 Study 5

In a study to validate an assessment centre utilised to select Female applicants for officer training in the military (Tziner, 1984), peer ratings, situational exercise performance scores, dimension ratings, an overall assessment rating (likelihood of success), and intelligence scores were used as predictors and validated against performance in training as well as against subsequent job performance. The training criterion was a weighted composite of scores obtained on objective examinations on the content of the training course, and instructors' ratings of performance on command exercises. The job performance criterion was a rating by commanding officers and their deputies. The correlations obtained are reflected in Table 4.4.

**TABLE 4.4**  
*Validity of Various Assessment Centre Predictors  
 with Training and Job Performance as Criteria*

Predictors	Criteria	
	Training performance	Job performance
Peer rating	0,51*	0,12
Situational exercise score	0,76**	0,18
OAR (likelihood of success)	0,60**	0,38*
Dimensions rating	0,58**	0,20
Verbal intelligence	0,17	0,14
Non-verbal intelligence	0,12	0,03
Inductive intelligence	0,08	0,07

(Tziner, 1984)

\*  $p < 0,05$

\*\*  $p < 0,01$

According to Tziner (1984), these results show that all the assessment centre measures are more superior predictors of officer training results than the traditional predictors. The author also points out that it is important to note that the situational exercises score has the highest validity of the assessment centre measures, showing that in this case, situational performance tests prove to be the most realistic testing ground for the demands of successful performance on officer training programmes. Tziner claims that a reason for these significant validities can be found in the fact that, compared to the study by Tziner and Dolan (1982), the content of this course was not academically oriented, but consisted of various practical field exercises. Thus,

intelligence scores played a minor role, as other skills and behaviour besides intelligence is needed to succeed in such exercises. As the situational exercises were in this case more related to the content of the course, they also provided the highest validities, confirming Asher and Sciarrino's (1974) point-to-point theory of the role that a closely related predictor-criterion plays in ensuring a high validity. According to Tziner (1984), the study further showed that if various predictors are combined statistically, even higher validities can be obtained regarding the prediction of managerial training results, which in this case produced a correlation coefficient of 0,83.

#### **4.2.6 Study 6 and 7**

Two studies were conducted to determine the validity of Police assessment centres (Extended Interviews: EIs) in Britain. In the first, Mays (1972), as quoted by Feltham (1988), found a significant but not very high relationship between the EI, OAR and potential achieved in Special Course training ( $r = 0,29$ ). In the second, a predictive validity study conducted by Feltham used various EI measures and an OAR as predictors, and training grades, supervisory ratings and rank attained as criteria. The training criteria consisted of a professional studies grade, an academic studies grade, and an overall training performance grade. Correlations of the EI measures with training criteria within one year of the EI are shown in **Table 4.5**.

**TABLE 4.5**  
*Correlations of Extended Interview (Assessment Centre) Measures with Training Criteria*

Extended interview measures	Training criteria		
	Academic studies	Professional studies	Overall training performance
Group discussion	-0,01	-0,01	-0,18*
Committee member	0,14	-0,02	0,22*
Written appreciation	0,11	0,05	0,01
General Information Test (GIT)	0,14	0,07	0,09
Peer nominations (Senior officer)	0,00	0,02	0,17*
Chairman interview	-0,01	-0,02	0,09
OAR	0,00	0,04	0,11

(Feltham, 1988)

\* $p < 0,05$

These results show that EI measures do not correlate significantly with academic and professional grades. Although three EI measures - group discussion ( $r = -0,18$ ), committee member ( $r = 0,22$ ) and peer nominations ( $r = 0,17$ ) - correlate significantly with overall training performance, the OAR does not ( $r = 0,11$ ). According to Feltham (1988), this study indicates that some EI components and combinations of EI components have validity for training criteria, but that the OAR does not have much validity as a predictor. In general, then, Feltham says that this specific Extended Interview (EI) is not a good predictor of Police training results, and suggests that a major reason for this could be due

to the fact that the EI procedure was not originally designed for this Police application, and that a further look will have to be taken at its design.

#### 4.3. DISCUSSION

All the above studies utilising managerial training results as criteria concerned military or Police officers, and Civil Service Administrators. From these studies as well as from the meta-analysis studies by Gaugler et al (1987), and Schmitt et al (1984), it can be seen that most private sector assessment centres use some or other job performance measures such as potential, career advancement, rank attained, supervisory ratings, and salary progression, as criteria in validity studies.

The writer feels that a possible reason for this could be that the private sector, being particularly concerned with obtaining a return on investment made in assessment activities, is more interested in predicting job performance than the public sector when justifying the use of assessment centres. A second reason could be that training performance is often used as a criterion because of its convenience - measures are readily available - while actual job performance is what is really of interest (Sackett & Ryan, 1985). A third reason could be the failure by some researchers (Klimoski & Strickland, 1977; Turnage & Muchinsky, 1984; Tziner & Dolan, 1982), to find support for the use of predictors such as supervisory evaluations, which may have consequently prompted the use of training rather than on-the-job measures as criteria. Whatever the

reason, Sackett and Ryan (1985) claim that any conclusions about the value of a particular assessment centre are likely to depend on the nature of the criterion measure used.

Of the studies reviewed above, those by Feltham (1988), Gardner and Williams (1973), and Mays (1972) produced the lowest validities, being 0,11, 0,21 and 0,29 respectively. The reason for the low validities obtained by Feltham (1988) and Mays (1972), who studied the same type of assessment centre to select Police recruits, is given by Feltham (1988, pp. 142-143) himself: "*The procedure was not originally designed for this police application.*" In fact, this specific Extended Interview (assessment centre) was originally developed for the selection of Civil Service Administrators. This supports the view of Thornton and Byham (1982), who stress the importance of developing a specific assessment centre for a specific purpose in a specific organisation.

The studies in which the highest validities were obtained by utilising managerial training results were those by Bornman (1982), Tziner (1984), and Tziner and Dolan (1982). These authors obtained correlation coefficients of 0,53, 0,60 and 0,38 respectively. In these studies, the criterion utilised was training success, and the assessment centres in question were designed specifically to measure dimensions that had been identified as important for success in training. This is emphasised by Tziner and Dolan (1982, p. 729) who state that during the job analysis phase of the development of the assessment centre, dimensions were identified which "*reflected capabilities,*

*qualities, and behaviors required for success as female officers in the military in training".* In these studies, it would appear that the significant validities obtained were due to what Wernimont and Campbell (1968) refer to as the consistency theory, as *behaviour* on the assessment centre was used to predict *behaviour* on a training course.

It is interesting to note that in some cases, intelligence tests are often better predictors of training success than assessment centre techniques (Tziner & Dolan, 1982), while in other cases not (Tziner, 1984). The explanation for these results appears to be that when the content of a course is academical or theoretical, intelligence tests are likely to be good predictors of training results. However, when the content of the training course involves the practical application of knowledge and skills in different situations, assessment centres, specifically designed to predict training results, or which use situational exercises designed to tap the dimensions essential for succeeding in training, are likely to be the best predictors of training results. This reasoning would also confirm Asher and Sciarrino's (1974) point-to-point theory regarding predictor-criterion similarity.

#### 4.4 SUMMARY

In this chapter the use of assessment centres as predictors of managerial training results was discussed. Previous studies in which training results were utilised to validate assessment centres were reviewed, and research results indicate that assessment centres specifically

designed to predict training results produce significant validities. The results also indicate that intelligence tests are valid predictors of training results when the content of a training course is academically or theoretically oriented.

This concludes the literature survey on the use and validity of various assessment instruments, including the assessment centre, to predict managerial as well as non-managerial training results. In the following chapters, evidence from an empirical study of an existing Middle Management Assessment Centre (MMAC) used as a selection device for entry to a Senior Management Training Course (SMTA) will be examined, to determine its validity in predicting managerial training results. The methodology followed for the study will first be explained, after which the research results obtained will be reported and discussed, in order to reach conclusions and make recommendations regarding the use of assessment centres for predicting managerial training results.

## CHAPTER 5

### METHOD OF INVESTIGATION

#### 5.1 INTRODUCTION

In this chapter, the methodology of the empirical study will be explained. First, a hypothesis will be formulated based on the completed literature survey. Secondly, the setting to the study will be given, after which the composition of the research sample will be discussed. Following this, all measuring instruments used in the study will be explained, after which the research design and procedure followed to collect the required research data will receive attention. The chapter will conclude with a discussion of the data-analysis methods used.

#### 5.2 HYPOTHESIS FORMULATION

Based on the literature survey in the previous chapters, and with the research question in mind, the following hypothesis can be formulated :

***THE ASSESSMENT CENTRE IS A PREDICTOR OF MANAGERIAL TRAINING RESULTS***

#### 5.3 SETTING

In order to test the stated hypothesis, a study was conducted within a service organisation which utilises a Middle Management Assessment Centre (MMAC) as a device to select middle managers for entry to a Senior Management Training Course (SMT). This assessment centre was initiated in 1989 and by February 1994, had completed its forty-

fifth session. By this date over 702 middle managers had participated in the programme.

#### **5.3.1 Purpose of the MMAC**

The purpose of the MMAC is to select middle managers for entry to a Senior Management Training Course (SMTc). Only those middle managers who obtain at least 60 percent on the assessment centre may attend the SMTc, which extends over a period of 11 months, representing a substantial financial investment by the organisation in each candidate. The aim of the SMTc is to provide middle managers with the knowledge and skills required to function on the next higher managerial level. Successful completion of the SMTc is one of the requirements for promotion to the next higher managerial rank.

#### **5.3.2 Participation**

Candidates in this service organisation are nominated by their functional heads to attend the MMAC. These nominated candidates are managers who are highest on their specific functional seniority list in terms of years of service, or managers who have been identified for a particular higher-level post and have to be given preference for attending the SMTc. Once nominated, it is compulsory for a manager to attend the MMAC, unless he has valid reasons for his non-attendance.

#### **5.3.3 Candidate acceptance for the Senior Management Training Course (SMTc)**

At the completion of an assessment centre, assessment data is consolidated by an administrator who uses predetermined weights for each dimension

to statistically calculate a weighted dimension rating. These weighted dimension ratings are summated and, using a predetermined formula, a weighted overall assessment rating (OAR) is calculated and expressed as a percentage. A candidate who obtains 60 percent or more is then nominated for the SMTc, while candidates scoring below this are not permitted to attend the course. These unsuccessful candidates are given the opportunity to attend specific dimension oriented training and are reassessed at a later stage to once again determine their acceptability for the SMTc.

No specific selection ratio is applied in selecting candidates for the SMTc, as all assessees who obtain 60 percent or more on the MMAC are granted permission to attend the course. One restriction is that a maximum of 40 candidates can be accommodated on each course, so successful MMAC assessees are prioritised according to their seniority in the organisation and nominated in this order. This often results in successful assessees having to stand over for a year or more before being able to attend a SMTc.

#### **5.4 SAMPLE**

To conduct a predictive validity study, predicting as well as criterion measures are required (Cascio, 1987). The sample for this study had thus to consist of subjects who had completed the MMAC, as well as the SMTc by January 1992, the date upon which this study was initiated. These inclusion requirements were necessary to ensure that a MMAC rating (predictor) as well as SMTc grades (criteria) were available for each subject.

During 1990 and 1991, seventy-seven middle managers were allowed to attend one of two SMTCs, having obtained an acceptable MMAC rating during 1989 and 1990. Eighteen managers who had obtained an OAR of less than 60 percent were not granted entry to the SMTTC. On completion of the second SMTTC at the end of 1991, training grades as well as an assessment centre rating (OAR) were available for all these 77 managers, fulfilling the abovementioned two requirements for inclusion in the research sample. These 77 managers were white males aged between 31 and 47 years, who had executive or supporting management exposure and with experience in the organisation ranging between 7 and 24 years. Table 5.1 gives more detail concerning the characteristics of the sample.

**TABLE 5.1**  
*Characteristics of the Research Sample*

Age range (years)	Number of subjects	Experience range (years)	Number of subjects
30 - 34,9	15	0 - 9,9	2
35 - 39,9	41	10 - 12,9	18
40 - 44,9	19	13 - 15,9	26
45 - 47,7	2	16 - 18,9	19
<u>Total sample:</u>	77	19 - 24,9	10
$\bar{x} = 37,5$ years		Unknown	2
		<u>Total sample:</u>	77
		$\bar{x} = 15$ years	

## 5.5 MEASURING INSTRUMENTS

### 5.5.1 The Middle Management Assessment Centre (MMAC)

#### (a) Development of the MMAC

The development of the MMAC was carried out in accordance with the steps and guide-lines provided by Thornton and Byham (1982), although various adaptations were made to make the MMAC organisation specific. These steps are outlined in **Appendix A**.

The job analysis that was conducted used the STAR interview method (Situation, Task, Action, Result), the critical incident technique (Flanagan, 1954) and the repertory grid technique (Stewart & Stewart, 1976). Twelve dimensions were derived, reflecting behaviour required for success in job positions on the next higher rank after middle management, and these dimensions were grouped into four categories as reflected in **Figure 5.1**.

A questionnaire was sent to a sample of current middle managers in the organisation, requesting them to rate the importance and frequency of use of each of the 12 formulated managerial dimensions. These ratings were then statistically analysed to determine a weight for each dimension regarding its importance. The final 12 dimensions and their weights were presented to top management who, after minor adjustments, approved them for use in the MMAC. These dimensions and weights are subsequently used to statistically calculate a weighted OAR for each candidate completing the MMAC.

<u>Dimensions</u>	<u>Categories</u>
Analytical ability Judgment Decisiveness	Decision-making
Planning and organising Delegating Control	Management
Group leadership Individual leadership Tenacity Negotiation skills	Interpersonal
Oral communication Written communication	Communication

**FIGURE 5.1 : Categories of MMAC dimensions**

Five situational exercises were designed to ensure adequate assessment of the identified dimensions, namely, an in-tray, a written analytical problem, two leaderless group discussions (LGD), and an individual interview which a candidate is required to conduct with a role player. The exercises were constructed on similar lines to those proposed by Crooks (1977), and Thornton and Byham (1982). Table 5.2 lists the dimensions assessed in each situational exercise.

**(b) Assessment staff**

The assessment staff consists of an administrator, a coordinator, assessors, and supporting clerical staff. Eight-member assessor teams are used during each assessment centre, which assesses 16 middle

**TABLE 5.2**  
*Dimensions Assessed in each Situational Exercise*

Dimensions	In-tray	Analysis problem	Group	Discussion	Interview
			Competitive	Consensus	
Analytical ability	X	X			X
Judgement	X	X			X
Decisiveness	X				X
Planning and organising	X	X			X
Delegating	X				
Control	X				X
Group leadership			X	X	
Individual leadership	X				X
Tenacity			X	X	X
Negotiation skills			X	X	X
Oral communication			X	X	X
Written communication	X	X			

managers. Half of the assessors are experienced middle or higher level managers and the other half are psychomotrists or psychologists, and each assessor pair, assigned to observe four specific candidates, consists of one manager and one psychomotrist or psychologist. Each assessor receives 3 days' training prior to an assessment centre, but if used again for successive assessment centres, is only required to attend a 1-day refresher course.

### (c) Assessment process

A MMAC is conducted over a period of  $3\frac{1}{2}$  days:

- \* Day 1. In-tray, analysis exercise and consensus group discussion.

- \* Day 2. Competitive group discussion and individual simulation interview.
- \* Day 3. Intergration of all assessment data by the administrator and assessors. At this joint meeting, the ratings obtained by each assessee on each dimension are pooled and discussed until a consensus rating for the dimension is agreed upon. This procedure is followed until all the dimensions for a specific assessee have been rated. Once an assessee has been discussed, the next assessee is discussed dimension-by-dimension in the same way. A detailed assessment report covering each dimension is then compiled by the assessors for each assessee. The assessors do not determine an OAR, which is done at a later stage by the administrator.
- \* Day 4. Each assessor gives individual feedback to two of the assessees he/she has observed, and only discusses performance on each dimension in terms of *satisfactory* or *less than satisfactory* behaviour. Ratings are not made known to the candidate nor is he/she advised whether or not he/she has been selected for the SMTc.

#### **5.5.2 Independent variables**

##### **(a) Consensus dimension ratings**

An assessment centre is designed to measure level of functioning on several separate dimensions of job success (Moses, 1987; Thornton & Byham, 1982; Klimoski & Brickner, 1987). These ratings are utilised, amongst others, by assessors to arrive at

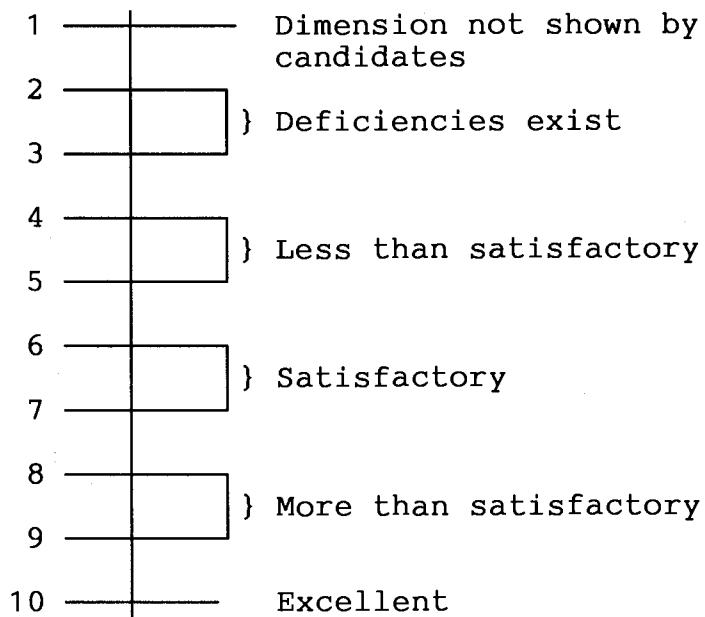
an OAR for each assessee (Moses, 1977; Sackett & Ryan, 1985; Gaugler & Thornton, 1989). While this final OAR has received the most attention in validity studies (Blanksby & Iles, 1990), few report on the predictive validity of the individual dimension ratings on which the OAR is based (Thornton & Byham, 1982). To address this identified research need, final individual dimension ratings are utilised as predicting measures in this study and correlated with training criteria.

During each of the group simulation exercises and the individual simulated interview exercise, candidates are observed by assessors who record their responses word for word. Following this, each dimension tapped by the specific exercise is identified and rated on a 10-point scale by the assessor. In the in-tray and analysis exercises, each candidate's product is marked and the applicable dimensions are identified and rated on the same 10-point scale. The scale used is given in Figure 5.2.

During data-integration, all assessors pool their ratings and, after discussion, a consensus rating is allocated to each dimension on the same 10-point scale. These 12 final consensus dimension ratings, expressed as a numerical measure out of a possible maximum of 10, are used as individual predicting measures in this study.

**(b) Average dimension score**

Assessors on the MMAC do not judgmentally determine an OAR by considering all relevant information available on each candidate. A weighted OAR is



**FIGURE 5.2 : Ten-point rating scale used to rate dimensions during each situational exercise**

statistically calculated by multiplying predetermined weights with each dimension rating. It is thus not possible to compare what Sawyer (1966) calls a clinical method of combining assessment data (judgmentally determined OAR not done in the MMAC) to a mechanical one, as numerous assessment centre researchers have done (Borman, 1982; Thornton & Byham, 1982; Tziner & Dolan, 1982; Tziner, 1984). If an average dimension rating is calculated as a "pure" mechanically determined OAR, being the mean of the dimension ratings, (Sackett & Ryan, 1985), it can then be compared to the statistically calculated "weighted" OAR. Comparing the predictive validities of these two measures allow inferences to be drawn regarding the validity of the weights allocated to the dimensions for calculating the weighted OAR. For this reason, an average dimension rating is utilised as a predictor in this study. As this rating is calculated by

adding all 12 consensus dimension ratings and dividing it by 12 to obtain a statistical average dimension score for each assessee, it can be regarded as a mechanically determined OAR for the MMAC.

**(c) Weighted Overall Assessment Rating (OAR)**

As the aim of the MMAC is to determine entry to the SMTC, prediction of training criteria is the fundamental issue (Gatewood & Feild, 1990), and as the OAR is the measure ultimately used to determine entry to the SMTC, using it as the main predicting measure in this study is based on what The Guidelines (1992) refer to as a logical foundation for predictor choice.

The use of the OAR is also based on what The Guidelines (1992) refer to as an empirical foundation, as a review study by Blanksby and Iles (1990) found that most of the research on assessment centre validity concentrated on the OAR as a predictor of managerial performance.

There is also a sound theoretical foundation for using the OAR to determine the predictive validity of the MMAC. This fact is based on a large volume of available theoretical literature on the design and implementation of assessment centres, which prescribes the determination of an OAR as the final measure of assessment centre performance (Jaffee, 1971; Keil, 1981; Moses, 1977; Thornton & Byham, 1982; Woodruffe, 1990).

The weighted OAR used by the MMAC is calculated by the administrator who multiplies each consensus dimension rating with its own specific numerical

weight (Section 5.5.1 (a) describes how these weights were determined). Each assessee's weighted dimension ratings are added together and a final percentage is calculated by using a maximum obtainable weighted total of 850 points as the denominator.

The final total weighted dimension rating, used as the OAR for each candidate, was utilised as the main predictor in this study. The calculation of this weighted OAR is explained in Table 5.3.

TABLE 5.3  
*Calculation of the MMAC Weighted OAR*

Dimension	Consensus rating	Fixed weight	Weighted rating
1	x	11	11x
2	y	6	6y
3	.	5	.
4	.	12	.
5	.	9	.
6	.	8	.
7	.	10	.
8	.	7	.
9	.	3	.
10	.	4	.
11	.	5	.
12	.	5	.
$\Sigma$			= z
OAR = [z ÷ 850 X 100 ÷ 1] %			

#### (d) Moderator variables

Cascio (1987) explains that in psychological research, differential predictability exists when the correlation between a predictor and criterion is found to vary as a function of classification on some third variable. Saunders (1956) calls such a

third variable a moderator, explaining that it could account for a portion of correlation variance.

Characteristics which have been investigated as potential assessment centre moderators include a candidate's age and departmental experience (Burroughs, Rollins & Hopkins, 1973) and sex (Moses & Boehm, 1975). Burroughs et al report a correlation of -0.34 ( $p < 0.01$ ) between the OAR and age among middle managers and found that managers from the plant department performed better than those from the engineering and commercial departments. Reviewing the research on sex and race, Thornton and Byham (1982) found inconclusive evidence on which generalisations could be made.

The sample for this study varied widely in terms of experience (years of service) and age (See Table 5.1). During assessment centres, assessors also noted that there appeared to be a difference in the performance of candidates from executive managerial positions in the organisation, compared to those from supporting managerial positions. Given these real (age and experience) and assumed (functional experience related performance) differences amongst the subjects, and the research results reported by Burroughs et al (1973) on the role of moderators on assessment centre validity, it was decided to investigate three MMAC variables as potential moderators, namely Age, Functional Exposure and Experience (years of service).

The potential moderator variables of Age and Experience were expressed as completed years, months less than 12 not being considered. Functional Exposure, being the managerial function

each subject had been exposed to, was measured with a 2-point numerical scale. A subject exposed to executive managerial functions was allocated a number of 1, and a subject exposed to supporting managerial functions 0. The numerical values for the three potential moderators were correlated with the three training criteria to determine their effect on MMAC validity.

**(e) Summary of independent variables**

Table 5.4 gives a summary of all the independent variables utilised in this study, as well as the nature of each measurement. Four variable groupings were used, namely, assessment center dimensions, average dimension score, weighted OAR and moderator variables.

**5.5.3 Dependent variables**

The dependent variables utilised in this study were the training grades which each trainee received at the completion of the SMTc.

**(a) The use of training grades as criteria of success**

In a meta-analysis of assessment center validity, Gaugler et al (1987) found that one of the five main categories of criteria most commonly used in validity studies is training results. According to Sackett and Ryan (1985), a reason for this is the readily availability of such measures. Supporting the use of training results as a criterion, Gatewood and Feild (1990) argue that such data is obtained under controlled conditions, while the legal permissibility of such data as a criterion measure is confirmed by The Guidelines (1992).

**TABLE 5.4**  
*Summary of Independent Variables*

Predictors and groupings	Nature of measurements
<u>Assessment centre dimensions</u>	
1 Analytical ability	Numerical rating on a scale of 1 to 10
2 Judgment	
3 Decisiveness	
4 Planning and organising	
5 Delegating	
6 Control	
7 Group leadership	
8 Individual leadership	
9 Tenacity	
10 Negotiation skills	
11 Oral communication	
12 Written communication	
<u>Average dimension score</u>	}
	Numerical rating as for dimensions
<u>Weighted OAR</u>	}
	Percentage score
<u>Moderator variables</u>	
1 Functional division	}
	Numerical rating of 1 or 0
2 Age	}
	Years
3 Previous experience	}
	Years

(b) **Multiple versus composite criteria**

Cascio (1987) states that although applied psychologists agree that performance in the work situation is multi-dimensional in nature, requiring multi-dimensional measurement, there is no agreement on how criteria should be handled. For him, the basic question is whether the various criterion measures should be combined into a composite score, or whether each should be treated separately. This view is supported by Schmidt and Kaplan (1971), stating that there has been a long controversy in industrial psychology about the issue of composite versus multiple criteria.

For Schmidt and Kaplan (1971), the resolution of the dilemma *will depend on the objectives of the researcher*. The authors explain that if the goal is increased psychological understanding of predictor-criterion relationships, then criterion elements are best kept separate. If managerial decision-making is the objective, then criterion elements should be weighted into a composite, representing a construct of overall worth to the organisation. Cascio's (1987) solution is that there should not be an insistence on the use of one *or* the other method of validation, but that both should be used for purposes of increased utility and understanding.

As improved organisational decision-making and increased predictor-criterion understanding are both important in this study, Cascio's (1987) suggestion to resolve the dilemma was followed, and a composite as well as multiple criteria were utilised.

(c) Measurement of multiple criteria used

Various managerial and operational principles and facts are taught on the SMTA, and trainees have to write numerous paper-and-pencil knowledge tests to pass these subjects. The grades obtained for these tests should, according to Kirkpatrick (1979), provide a valid measure of the amount of theoretical learning that has taken place. This measure, referred to as a Theoretical Grade, was utilised as one of the separate training criteria. It consists of an average percentage score, calculated for each trainee, by using all the theoretical knowledge test scores which he obtained during the SMTA. These tests covered various functional and general management subjects and were written at various stages during the 11 month course.

During the problem-solving exercises presented on the SMTA, trainees have to apply the theoretical knowledge gained to solve practical managerial and operational problems. A trainee explains his/her reasoning and suggested solution in writing, for which a practical grade is allocated, and this determines whether he/she passes the practical phase or not. As these exercises assess what a trainee *is able to do*, they conform to what Gatewood and Feild (1990) call performance tests. The grades obtained for these tests should, according to Kirkpatrick (1979), provide a valid measure of the amount of practical learning that has taken place. This measure, referred to as a Practical Grade, was utilised as the second separate training criterion. It consists of an average percentage mark, calculated for each trainee by using all practical problem-solving

exercise scores which he/she obtained during the SMTA. These exercises comprised management and operational problems, which had to either be solved individually or as a team.

**(d) Measurement of the composite criterion used**

When a composite criterion is used, Nagle (1953) argues that it should provide a yardstick or overall measure of success or value to the organisation for each individual. He claims that such a single index is indispensable in decision-making and individual comparisons.

At the completion of the SMTA, it must be decided who has successfully completed the SMTA or not. The overall index on which this decision is based is a Total Course Grade, for which a trainee must obtain at least 60 percent to pass the course. The Total Course Grade is obtained by combining each trainee's theoretical and practical grades and calculating an average measure. This measure, referred to as a Total Grade, complies with Nagle's (1953) required overall index of success, and was utilised as the composite training criterion in this validity study. It thus consists of both theoretical and practical training components.

**5.6 RESEARCH DESIGN AND PROCEDURE**

**5.6.1 Assessment centres and predictive validity**

When an assessment centre is used as a selection technique, Thornton and Byham (1982) state that its effectiveness needs to be demonstrated by evidence of predictive validity. This view is supported by Sackett (1987), who states that centres being used

to specifically select individuals for entry into any training programme, should be validated by means of the criterion-related validity model. As the MMAC is used to select managers for the SMTC, the reasoning by Thornton and Byham (1982) and Sackett (1987) requires that its predictive validity be determined by correlating predictor and criteria measures with each other.

#### **5.6.2 Obtaining predictor and criterion data**

All the data required for this study was available on the subjects' personal files, and authority was required, and obtained, to gain access to them. The information regarding each candidate's performance on the MMAC (predicting measures) was obtained from a standardised assessment centre exercise-dimension matrix which had been compiled for each assessee by the assessment centre administrator. An example of such a matrix is attached as **Appendix B**. The information regarding functional exposure, age and experience (moderator variables), was obtained from available computer records for each subject. The SMTC grades for each subject (criterion measures) were obtained from the Final Course Report, compiled for each SMTC trainee at the completion of the course.

#### **5.7 DATA-ANALYSIS METHODS**

According to Kerlinger (1973), behavioural problems are almost all multivariate in nature and cannot be solved by a bivariate approach which considers only one independent and one dependent variable at a time. When multiple independent variables must be correlated with multiple dependent variables, Kerlinger states that multivariate analysis must be

used. In the present study, consisting of multiple assessment measures and multiple training criteria, multivariate analysis techniques were used to analyse the data. The independent and dependent variables were coded for statistical processing by means of the Statistical Package for the Social Sciences (SPSS) computer programme (SPSS User's Guide, 1983), and the specific analysis techniques used are discussed below.

#### **5.7.1 Intercorrelations between individual dimension ratings, average dimension rating and OAR**

According to Turnage and Muchinsky (1984), it is important to determine how the various assessment centre ratings and measures relate to each other by computing their statistical intercorrelations. Thornton and Byham (1982) state that these intercorrelations are essential to provide research evidence about which dimensions contribute the most to the determination of the OAR, as this could lead to revisions in exercises, dimensions and assessor training. In the case of the MMAC, in which dimensions are judgmentally determined while the OAR is a statistical measure calculated by using predetermined dimension weights, insight can be gained into the validity of these weights by intercorrelating the dimension ratings with the weighted OAR and the average dimension rating. As the dimension ratings, average dimension score and OAR are all continuous variables, Pearson's correlation coefficient as prescribed by Kerlinger (1973) was used to provide a measure of the degree of the relationship between paired values of the various measures.

### **5.7.2 Intercorrelations between training criteria scores**

Cascio (1987) states that it is important to determine the equivalence of criteria by correlating them with each other. He explains that if criteria correlate perfectly, they are equivalent and measure the same elements, with implications for a validity study. As multiple training criteria are utilised to determine the predictive validity of the MMAC, their equivalence will have to be determined to obtain insight into the operation of the various assessment centre measures used as predictors. As the training criterion measures, being percentage marks, are all continuous variables, Pearson's correlation coefficient as prescribed by Kerlinger (1973) was used to determine the degree of intercorrelation between them.

### **5.7.3 Validity of predicting measures with training results as the criterion of success**

In order to determine the predictive validity of a selection device, it will have to be shown that a relationship exists between the predicting measures and the criterion of success (Cascio, 1987). As the MMAC is used as a selection device to determine entry to the SMTc, the OAR on which this entry decision is based will have to be correlated with the three training criteria, in order to determine its predictive validity. As the OAR and training grades utilised as criteria are all continuous variables, Pearson's correlation coefficient was used to determine the degree of relationship between them (Kerlinger, 1973).

To provide additional information on the predictive validity of individual dimension ratings compared to the OAR, a research need identified by Thornton and Byham (1982), final dimension ratings were correlated with the three training criteria. Individual dimension validities also provide information on the importance of individual dimensions in predicting outcome criteria, and, as described by Thornton and Byham, contribute to determining reliable weights for use in calculating an OAR for use in making practical decisions. As the final dimension ratings and the training grades utilised as criteria are all continuous variables, Pearson's correlation coefficient was used to determine the degree of the relationship between paired values of these variables (Kerlinger, 1973).

In order to compare the predictive validities of the statistically determined average dimension rating (a "pure" mechanical method of combining assessment data), and the weighted OAR, the average dimension rating was correlated with the three training criteria. This comparison allows for inferences to be made concerning the validity of the predetermined weights used to calculate the weighted OAR. As the average dimension rating and the training grades utilised as criteria are all continuous variables, Pearson's correlation coefficient was used to determine this validity coefficient.

#### **5.7.4 The contribution of moderator variables to assessment centre validity**

According to Gaugler et al (1987), individual investigations of moderators, using single samples, have yielded conflicting and inconclusive evidence.

In order to investigate the role that moderator variables play in the validity of the MMAC, candidates' age, functional exposure in the organisation and experience were correlated with results achieved on the SMTA. As age, experience and the training grades utilised as criteria are all continuous variables, Pearson's correlation coefficient was used to calculate the validity coefficients (Kerlinger, 1973). As the functional exposure score is an artificial dichotomy, consisting of nominal values of 1 or 0, a biserial correlation coefficient was computed between paired values of these nominal measures and the training grades utilised as criteria, which are continuous variables (Flowers, 1981). This provided correlation coefficients comparable to the validities calculated for the OAR, individual dimension ratings and the average dimension rating.

#### **5.7.5 Multiple linear regression analysis**

Various researchers have investigated the increase obtained in assessment centre prediction by using combinations of predicting measures and computing multiple correlation coefficients for the various combinations (Huck, 1973; Jones, Herriot, Long and Drakely, 1991; Tziner & Dolan, 1982; Wollowick & McNamara, 1969). In order to investigate prediction accuracy with regard to the MMAC, independent variables were classified into four types of predicting measures, namely Dimension Ratings, OAR, Age and Experience. Multiple correlation coefficients were computed by conducting a linear regression analysis of SMTA training grades on these four predicting measures, as described by Weiss (1976). These multiple correlations were utilised to determine the individual and combined

accuracy of the Dimension Ratings, OAR, Age and Experience measures in predicting training results achieved on the SMTC, as well as to determine the portion of criterion variance that each predictor type explained.

#### 5.7.6 Multiple stepwise regression analysis

Given the cost and time involved in operating an assessment centre (Cascio & Silby, 1979; Thornton & Byham, 1982), numerous assessment centre researchers have investigated the minimum number of assessment predictors which make a maximum contribution to prediction of managerial performance (Tziner, 1984; Tziner & Dolan, 1982; Thornton & Byham, 1982; Wollowick and McNamara, 1969). According to Wollowick and McNamara, such research can provide information on how many assessment instruments are necessary for maximum validity, and which individual measures can be eliminated without significantly reducing the predictive validity of the total assessment process. This is supported by Turnage and Muchinsky (1984), who argue that examining the incremental predictive validity value of various measures becomes imperative, as costs could be reduced by using only the most valid predictors.

To determine the relative contribution to overall validity of each type of MMAC predictor measure, a multiple stepwise regression analysis was conducted, using the same four types of predictor measures utilised in the linear regression analysis, namely the Dimension Ratings, OAR, Age and Experience. According to Kerlinger (1973), the SPSS computer programme stepwise procedure entails selecting the predictor measure which has the

highest correlation with the training criterion, and calculating regression statistics. It will then select the next predictor measure which, after the first measure, will contribute most of the variance of the training criterion. It will then examine the contribution the first measure will have made, had it been entered second. If this contribution turns out to be not statistically significant, the measure will then be dropped. The process is continued until a statistical test of significance strikes a predictor measure that does not contribute significantly to the training criteria.

#### **5.7.7 Analysis and interpretation of results**

The research results obtained by means of the statistical computer processing were analysed and interpreted in order to finally reach conclusions and make useful recommendations regarding the validity and use of the Middle Management Assessment Centre as a selection device for the Senior Management Training Course.

#### **5.8 SUMMARY**

In this chapter, the methodology for this study was explained. First, a hypothesis was formulated, after which the setting to the study was given. The composition of the sample was discussed and shown to be 77 white male middle managers, and all measuring instruments utilised in the study were explained. The research design and procedure followed to collect the required data were explained, and the chapter was concluded with a discussion of the data-analysis methods which were used.

In the next chapter, the research results obtained from each statistical analysis will be reported. Intercorrelations within the various variables, correlations between predicting measures and training criteria, the role of moderator variables and the results obtained from the multiple linear and stepwise regression analyses will be given.

**CHAPTER 6****RESEARCH RESULTS****6.1 INTRODUCTION**

In this chapter, the results obtained from the statistical analyses will be reported. First, the intercorrelations obtained between the various variables will be given. Following this, the correlations obtained between the Middle Management Assessment Centre (MMAC) predicting measures and the Senior Management Training Course (SMTc) grades will be reported, to indicate the validity of the MMAC as a predictor of managerial training results. The contribution of various moderator variables to the validity of the MMAC will also be shown, and the results of the multiple linear and stepwise regression analyses that were conducted will be given.

**6.2 INTERCORRELATIONS BETWEEN VARIABLES****6.2.1 Intercorrelations between predictors**

The first statistical data produced was Pearson's correlation coefficients between individual dimension ratings, the average dimension rating and the OAR, with results reported in Table 6.1. All the dimensions are significantly intercorrelated with the OAR at the 0,01 level of significance, ranging from a minimum for Written Communication ( $r = 0,34$ ) to a maximum for Judgment ( $r = 0,82$ ). The fact that the Average Dimension Rating and OAR are highly intercorrelated ( $r = 0,99$ ) indicates that they both contain virtually the same conceptual space (Cascio, 1987), which is to be expected, as both are statistically calculated based on the same dimensions.

TABLE 6.1

Matrix of Intercorrelations between Individual Dimension Ratings, Average Dimension Rating and OAR

Assessment measures	1	2	3	4	5	6	7	8	9	10	11	12	13	$\bar{x}$	$s^2$
1. Analytical ability	-													72,8	9,7
2. Judgment	0,73**	-												68,7	10,2
3. Decisiveness	0,47**	0,46**	-											75,1	12,5
4. Planning and org	0,67**	0,72**	0,49**	-										71,9	9,7
5. Delegation	0,10	0,30**	0,27**	0,34**	-									83,7	11,8
6. Control	0,44**	0,54**	0,53**	0,49**	0,23*	-								69,6	14,6
7. Group leadership	0,36**	0,42**	0,33**	0,31**	0,13	0,27*	-							74,1	10,2
8. Indiv leadership	0,40**	0,64**	0,52**	0,52**	0,31**	0,53**	0,53**	-						72,5	15,8
9. Tenacity	0,46**	0,56**	0,59**	0,46**	0,12	0,55**	0,62**	0,56**	-					74,7	9,9
10.Negotiation	0,34**	0,53**	0,50**	0,52**	0,28*	0,42**	0,66**	0,75**	0,64**	-				72,5	11,0
11.Oral commu	0,34**	0,47**	0,39**	0,43**	0,20	0,41**	0,69**	0,54**	0,70**	0,74**	-			76,2	9,0
12.Written commu	0,43**	0,35**	0,20**	0,38**	0,07	0,25*	0,05	0,12	0,17	0,05	0,10	-		70,1	8,0
13.Ave of dimensions	0,68**	0,81**	0,72**	0,76**	0,42**	0,71**	0,64**	0,81**	0,78**	0,79**	0,71**	0,35**	-	73,5	7,6
14.OAR	0,70**	0,82**	0,68**	0,80**	0,45**	0,71**	0,63**	0,80**	0,74**	0,75**	0,68**	0,34**	0,99**	73,5	7,6

\* p < .05, two-tailed  
n = 77

\*\* p < .01, two-tailed  
n = 77

Except for Delegation ( $r = 0,45$ ) and Written Communication ( $r = 0,34$ ), all other dimension ratings intercorrelate highly with the OAR. This suggests that they are all measuring the same constructs. The two dimensions producing the lowest correlation with the OAR, namely Delegation and Written Communication, intercorrelate relatively low with all the other dimensions. This suggests that these two dimensions are measuring a specific ability not picked up by the others (Dulewicz, Fletcher & Wood, 1983). Another explanation could be that Delegation is only measured in one exercise (In-tray) and Written Communication in two (In-tray and Analysis Exercise). Assessees may therefore not have had ample opportunity to demonstrate the dimensions to the observers.

#### 6.2.2 Intercorrelations between criteria

The second statistical data produced was Pearson's correlation coefficients between the three training criteria, namely the Theoretical Grade, Practical Grade, and Total Grade. The results obtained are given in Table 6.2. Both the Theoretical Grade ( $r = 0,76$ ) and the Practical Grade ( $r = 0,87$ ) are significantly intercorrelated with the Total Grade, which indicates that both are measuring aspects similar to the Total Grade. This is to be expected, as the Total Grade is an average of the Theoretical and Practical scores. The Theoretical and Practical Grades intercorrelate relatively low ( $r = 0,36$ ), which suggests that they do not measure the same content of the training programme. This is also to be expected, as the Theoretical Grade measures theoretical course content and the Practical Grade practical content. Choosing them as two separate criteria is thus justified, each contributing some-

**TABLE 6.2**  
*Matrix of Intercorrelations between Theoretical, Practical, and Total Training Grades (Criteria)*

Criteria	1	2	3	$\bar{x}$	$s^2$	n
Theoretical grade	-	-	-	68,2	2,8	77
Practical grade	0,36**	-	-	66,1	3,9	77
Total grade	0,76**	0,86**	-	67,2	2,8	77

\*\* $p < 0,01$ , two-sided

thing unique to our understanding of the predicting measures (Cascio, 1987).

Important to note is the small variance in the Theoretical, Practical and Total Grade scores, namely 2,8, 3,9 and 2,8 respectively. According to Dobson (1988), the correlation between two variables is affected by their range, and a correlation used to estimate the validity of a measure will underestimate the true population validity if based upon a sample in which criterion scores are restricted in range. In this study, the restriction in range of training grades could produce what Schmidt et al (1976) refer to as restricted validities between MMAC predicting measures and these criterion measures.

### 6.3 CORRELATIONS BETWEEN PREDICTING MEASURES AND TRAINING CRITERIA

Table 6.3 presents validity coefficients between the predicting measures used in the MMAC, consisting of individual dimension ratings, an average dimension rating and the weighted OAR, and the criterion of

**TABLE 6.3**  
*Correlations between Assessment Centre Predictors  
 and Training Results (Criteria)*

Predicting measures	Training criteria		
	Theoretical grade	Practical grade	Total grade
<b><u>Individual Dimensions</u></b>			
1. Analytical ability	0,02	0,23*	0,18
2. Judgment	0,01	0,08	0,06
3. Decisiveness	-0,06	0,15	0,09
4. Planning and organising	-0,07	0,18	0,10
5. Delegating	0,05	0,09	0,10
6. Control	-0,13	0,05	-0,03
7. Group leadership	0,17	0,25*	0,27*
8. Individual leadership	-0,07	0,10	0,04
9. Tenacity	-0,00	0,21	0,15
10. Negotiation skills	0,02	0,05	0,05
11. Oral communication	0,06	0,07	0,09
12. Written communication	-0,08	0,05	0,00
<b>Ave dimension rating</b>	-0,02	0,18	0,13
<b>Weighted OAR</b>	-0,02	0,21	0,14

\* $p < 0,05$ , two-sided

interest, being three training measures obtained on the SMTG (Theoretical Grade, Practical Grade, and Total Grade).

#### 6.3.1 Validity of individual dimension ratings

No individual dimension rating correlates significantly with the Theoretical Grade as a criterion of training success. With the Practical Grade as the criterion, Analytical Ability ( $r = 0,23$ ) and Group Leadership ( $r = 0,25$ ) are the only dimensions which correlate significantly with the Practical Grade criterion ( $p < 0,05$ ). This can be expected, as Analytical Ability is required to apply

the theoretical knowledge gained on the SMTG to practical managerial and operational problems which have to be solved during the course, at times individually and often as a team. When the practical problems have to be solved as a team, successful team functioning and correct solutions are dependent on Group Leadership, which could explain the significant correlation obtained between this dimension and the Practical Grade score.

The only dimension which correlates significantly with the Total Grade is Group Leadership ( $r = 0,27$ ,  $p < 0,05$ ). As the Total Grade is an average of the Theoretical and Practical Grades, it can be expected that the dimensions which correlate significantly with the Theoretical Grade and Practical Grades, would also correlate significantly with the Total Grade. However, only Group Leadership correlates significantly with the Total Grade ( $r = 0,27$ ,  $p < 0,05$ ) and not Analytical Ability ( $r = 0,18$ ). This can be explained by analysing the intercorrelations between the training criteria (See Table 6.2). From this table it can be seen that the Practical Grade ( $r = 0,87$ ,  $p < 0,01$ ) correlates higher than the Theoretical Grade ( $r = 0,76$ ,  $p < 0,01$ ) with the Total Grade, indicating that the Practical and Total Grades measure more aspects in common than the Theoretical and Total Grades. Thus, only Group Leadership correlates significantly with both the Practical and Total Grades.

#### 6.3.2 Validity of average dimension rating

As explained, the average dimension rating is the arithmetic mean of all the dimension ratings, referred to by Thornton and Byham (1982) as a mechanically determined OAR. This rating produced

insignificant correlations of -0,02, 0,18 and 0,13 with the three training criteria respectively.

### **6.3.3 Validity of the weighted OAR**

From Table 6.3 it can be seen that the weighted OAR produced insignificant validities with the three training criteria. These results indicate that by allocating weights to the final consensus dimension ratings for a weighted OAR, predictive validity changed from -0,02, 0,18 and 0,13 for the average dimension rating against the three training criteria, to -0,02, 0,21 and 0,14 for the weighted OAR against the same three training criteria, all correlations being statistically insignificant. The use of dimension weights did thus not significantly improve the correlation coefficients obtained with the average dimension rating.

### **6.4 CONTRIBUTION OF MODERATOR VARIABLES TO ASSESSMENT CENTRE VALIDITY**

Validity coefficients of Functional Exposure, Age, and Experience with the three training criteria are presented in Table 6.4. From the results it can be seen that of the three potential moderator variables, only Age correlates significantly with all three of the training criteria, producing correlation coefficients of -0,33, -0,23, and -0,33. As these correlations are negative, it indicates that the older a candidate, the more likely he is to perform poorly on the SMTc. Age can thus be considered to be a significant moderator of MMAC validity. Regarding the low correlations obtained for Functional Exposure and Experience, it would seem that when considered on their own, these two variables do not play a significant moderating role in the validity of the MMAC.

**TABLE 6.4**  
*Validity coefficients of Functional Exposure, Age and Experience with Training Results as Criteria*

Predictor measures	Training criteria			$\bar{x}$	$s^2$	n
	Theory grade	Practical grade	Total grade			
Functional Exposure	0,09	0,13	0,13	0,5	0,5	74
Age	-0,33**	-0,23*	-0,33**	37,5	3,9	77
Experience	0,18	-0,10	0,03	14,9	3,5	75

\*  $p < 0,05$ , two-sided

\*\*  $p < 0,01$ , two-sided

## 6.5 INDIVIDUAL AND COMBINED ACCURACY OF VARIOUS MMAC PREDICTOR COMBINATIONS IN PREDICTING TRAINING RESULTS

The results of the multiple linear regression analysis of training grades on the predicting measures of Dimension Ratings, OAR, Age and Experience are reflected in Table 6.5.

### 6.5.1 Results with Theoretical Grade as the dependent variable

The combination of OAR, Age and Experience is the most statistically significant predictor of Theoretical Training Grades ( $R = 0,42$ ,  $p < 0,01$ ), explaining 18 percent of criterion variance ( $R^2 = 0,18$ ,  $p < 0,01$ ). When the Dimension Ratings are combined with Age and Experience, the multiple correlation coefficient increases to 0,51 and the criterion variance accounted for increases to 26 percent ( $R^2 = 0,26$ ), although both are statistically insignificant ( $p = 0,13$ ).

### 6.5.2 Results with Practical Grade as the dependent variable

The combination of OAR, Age and Experience is the most statistically significant predictor of Practical Training Grades ( $R = 0,36$ ,  $p < 0,05$ ), explaining 12 percent of criterion variance ( $R^2 = 0,12$ ,  $p < 0,05$ ). When the Dimension Ratings are combined with Age and Experience, the multiple correlation coefficient increases to 0,53 and the criterion variance accounted for increases to 28 percent ( $R^2 = 0,28$ ), although both are statistically insignificant ( $p = 0,08$ ).

### 6.5.3 Results with Total Grade as the dependent variable

The combination of OAR, Age and Experience is the most statistically significant predictor of Total Training Grades ( $R = 0,41$ ,  $p < 0,01$ ), explaining 17 percent of criterion variance ( $R^2 = 0,17$ ,  $p < 0,01$ ). When the Dimension Ratings are combined with Age and Experience, the multiple correlation coefficient increases to 0,52 and the criterion variance accounted for increases to 27 percent ( $R^2 = 0,27$ ), although both are statistically insignificant ( $p = 0,09$ ).

## 6.6 CONTRIBUTION OF OAR, AGE AND EXPERIENCE TO MMAC VALIDITY

From the multiple linear regression analysis, only the combination of OAR, Age and Experience is statistically significant in accounting for a portion of criterion variance (See Table 6.5). A stepwise regression analysis was consequently conducted on the three training criteria using this predictor combination to determine the unique con-

TABLE 6.5

Contribution of Various Predictor Combinations to Assessment Centre Validity

Source of variation	Mean square	F-ratio	Probability	Error of prediction	R	R <sup>2</sup>
<b>DEPENDENT VARIABLE : THEORETICAL GRADE</b>						
Dimension ratings	5,413	0,615	0,8215	554,367	0,324	0,1049
OAR	0,336	0,040	0,8417	618,986	0,022	0,0005
Dimensions, age and experience	11,595	1,517	0,1334	450,995	0,514	0,2647
OAR, age and experience	37,209	5,192	0,0027	501,699	0,427	0,1820
<b>DEPENDENT VARIABLE : PRACTICAL GRADE</b>						
Dimension ratings	21,543	1,501	0,1478	904,363	0,471	0,2223
OAR	49,191	3,269	0,0747	1113,684	0,206	0,0423
Dimensions, age and experience	22,938	1,665	0,0884	812,663	0,532	0,2832
OAR, age and experience	49,001	3,476	0,0204	986,787	0,360	0,1297
<b>DEPENDENT VARIABLE : TOTAL GRADE</b>						
Dimension ratings	8,250	1,043	0,4226	498,386	0,407	0,1657
OAR	11,631	1,469	0,2293	585,754	0,140	0,0195
Dimensions, age and experience	11,762	1,631	0,0973	425,461	0,528	0,2790
OAR, age and experience	34,118	4,896	0,0038	487,800	0,416	0,1734

tribution to criterion variance of the three individual predictor measures.

#### 6.6.1 Unforced stepwise regression analysis

Table 6.6 shows the results when the three predictors were entered stepwise without being forced. Considered alone, Age is the most accurate predictor of Theoretical Training Grades ( $R = 0,32$ ,  $p < 0,01$ ) and accounts for most of the criterion variance, namely 10 percent ( $R^2 = 0,10$ ). When Experience is added to Age, the multiple correlation coefficient increases to 0,42 ( $p < 0,01$ ), while explained variance increases to 18 percent ( $R^2 = 0,18$ ). The OAR did not make a significant contribution to variance, and was excluded.

With Practical Training Grades, no predictor on its own significantly explains any criterion variance. Age, considered alone, is the most accurate predictor, ( $R = 0,22$ ,  $p = 0,06$ ), and accounts for most of the criterion variance, namely 5 percent ( $R^2 = 0,05$ ). When the OAR is combined with Age, the multiple correlation coefficient increases to 0,36 ( $p < 0,01$ ), while the explained criterion variance increases to 13 percent ( $R^2 = 0,13$ ,  $p < 0,01$ ). Experience did not make a significant contribution and was excluded.

With the Total Training Grade, Age is the variable, considered alone, which is the most accurate predictor ( $R = 0,32$ ,  $p < 0,01$ ), and accounts for most of criterion variance, namely 10 percent ( $R^2 = 0,10$ ,  $p < 0,01$ ). When the OAR is combined with Age, the multiple correlation coefficient increases to 0,40 ( $p < 0,01$ ), while explained criterion variance increases to 16 percent ( $R^2 = 0,16$ ,  $p < 0,01$ ). Experience was once again excluded from the calculation.

**TABLE 6.6**  
**Stepwise Regression Analysis: OAR, Age and Experience (Unforced)**

Predicting measures	R	R <sup>2</sup>	Probability
Dependent variable: Theoretical grade			
<u>Step 1.</u> Age	0,32	0,10	0,005
<u>Step 2.</u> Age and Experience	0,42	0,18	0,001
OAR did not make a significant ( $p \leq 0,15$ ) contribution to variance and was excluded.			
Dependent variable: Practical grade			
<u>Step 1.</u> Age	0,22	0,05	0,060
<u>Step 2.</u> Age and OAR	0,36	0,13	0,008
Experience did not make a significant ( $p \leq 0,15$ ) contribution to variance and was excluded.			
Dependent variable: Total grade			
<u>Step 1.</u> Age	0,32	0,10	0,005
<u>Step 2.</u> Age and OAR	0,40	0,16	0,001
Experience did not make a significant ( $p \leq 0,15$ ) contribution to variance and was excluded.			

From the above stepwise regression analysis, the following can be noted regarding the three predictors:

- \* Considered individually, Age is the most accurate and only statistically significant predictor of all three training criteria, and explains the most criterion variance regarding the Theoretical and Total Training Grades ( $R = 0,01$ ,  $p < 0,01$ ).
- \* A combination of Age and Experience is the most accurate predictor of Theoretical Training Grades obtained on the SMTG ( $R = 0,42$ ,  $p < 0,01$ ) and explains the most criterion variance, namely 18 percent.

- \* A combination of Age and the OAR is the most accurate predictor of Practical Training Grades obtained on the SMTC ( $R = 0,36$ ,  $p < 0,01$ ) and explains the most criterion variance, namely 13 percent.
- \* A combination of Age and the OAR is the most accurate predictor of the Total Training Grade obtained on the SMTC ( $R = 0,40$ ,  $p < 0,01$ ) and explains the most criterion variance, namely 16 percent.

#### 6.6.2 Forced stepwise regression analysis

In order to ensure that the predictor which was excluded by the unforced stepwise procedure for each training criterion is unable to make a significant contribution to prediction accuracy and explain a larger percentage of criterion variance, the three individual predictors were forced into the stepwise procedure for each training criterion. Table 6.7 shows the results obtained.

TABLE 6.7  
Stepwise Regression Analysis: OAR, Age and Experience (All Forced)

	Theoretical grade		Practical grade		Total grade	
Predicting measure	R	$R^2$	R	$R^2$	R	$R^2$
OAR, Age and Experience	0,42**	0,18**	0,36*	0,13*	0,41**	0,17**

\*  $p < 0,05$

\*\*  $p < 0,01$

In the case of the Total Training Grade, a statistically significant increase in predictive accuracy was gained by forcing Experience into the procedure, as the multiple correlation coefficient

then increased from 0,40 to 0,41 ( $p < 0,01$ ). A statistically significant increase in the amount of criterion variance explained was also obtained, namely, from 16 to 17 percent ( $p < 0,01$ ). Although these increases are statistically significant, they are what Tziner and Dolan (1982, p. 734) refer to as "*marginal statistically speaking*". In the case of the Theoretical and Practical Training Grades, no statistically significant increase in predictive accuracy or variance accountability was obtained by forcing the excluded predictor into the stepwise procedure.

As the Total Grade achieved on the SMTG determines whether a trainee passes the course or not, the MMAC measure which is the most accurate predictor of this grade is a combination of each assessee's final weighted OAR, his age and his experience, producing a multiple correlation coefficient of 0,41 and explaining 17 percent of the training criteria variance ( $p < 0,01$ ).

#### 6.7 SUMMARY

In this chapter, the results obtained from the statistical analyses were reported. The first data shown was intercorrelations between all the variables. This was followed by a discussion of the correlations obtained between the various MMAC predicting measures and the training results achieved on the SMTG. The contribution of various potential moderator variables was reported, and age was identified as the most significant moderator of MMAC validity. Finally, the results of the multiple linear and stepwise regression analyses conducted to investigate the unique contribution of various combinations of predictors to MMAC validity were reported.

In the following and final chapter of this report, the research results reported in this chapter will be interpreted and discussed. The chapter will begin with a discussion of the predictive validity of the MMAC, after which the limitations of the study will be given. Conclusions reached and recommendations made will then be reported, and the chapter will conclude by giving directions for possible future research on assessment centres.

## CHAPTER 7

### DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

#### 7.1 INTRODUCTION

In this final chapter of the report, the research results reported in the previous chapter will be interpreted and discussed. The chapter will begin with a discussion of the predictive validity of the Middle Management Assessment Centre (MMAC), after which the limitations of the present study will be given. Conclusions reached and recommendations made will then be reported, and the chapter will conclude by giving directions for possible future research on assessment centres.

#### 7.2 DISCUSSION OF RESEARCH RESULTS

##### 7.2.1 Predictive validity of the MMAC

The present study investigated the validity of the MMAC as a selection device to determine entry to the Senior Management Training Course (SMT), utilising training performance as the criterion of success. The results of this study, when compared to the results of a meta-analysis of assessment centre validity conducted by Gaugler et al (1987), would seem to indicate that the MMAC does not have predictive validity. This comparison is reported in **Table 7.1**.

There are three possible explanations for this apparent lack of predictive validity of the MMAC :

**TABLE 7.1**

*Comparison of MMAC Validity to Results of Meta-Analysis of Assessment Centre Validity Conducted by Gaugler et al (1987)*

Sample	Mean validity	Lower bound	MMAC OAR
Total for 50 assessment centre studies	0,37	0,21	0,14
Studies utilising training as the criterion of success	0,35	0,17	0,14
Studies with assess- ment centres being used for selection purposes	0,41	0,34	0,14

Note. Mean validity = mean validity corrected for statistical artifacts; Lower bound = lower bound of the 90 percent credibility value for the average corrected validity coefficient; MMAC OAR = validity coefficient between MMAC OAR and Total Training Grade achieved on the SMTG

(a) Lack of point-to-point correspondence between predictor and criterion

According to the consistency theory of Wernimont and Campbell (1968), and the point-to-point theory of Asher and Sciarrino (1974), the highest validities will be obtained when the predictor and criterion are similar in nature. Wernimont and Campbell warn that for the consistency notion to be consistent, the measures to be predicted must also be measures of behaviour. During the MMAC, various managerial behavioural dimensions are assessed, which are used to calculate a weighted OAR. This OAR is used to predict training performance on the SMTG. However, the grades allocated to trainees during the SMTG, which determine success on the course, consist of various theoretical and practical

marks, and do not represent a direct measure of a trainee's managerial behaviour on the SMTA. The managerial dimensions on which the MMAC is based were not measured directly on the SMTA used in this study, and their validity as predictors of success on the course had to be inferred from the correlations obtained between the OAR and the theoretical and practical grades obtained on the SMTA. There is thus a lack of point-to-point correspondence between the MMAC predictors (behavioural dimensions) and the SMTA training performance criteria (theoretical and practical grades).

In order to test this explanation, it would be necessary to assess trainees on the SMTA according to specific behavioural dimensions representing successful trainee behaviour. These could be the same dimensions assessed during the MMAC (if they do in fact represent successful trainee behaviour), or a set of dimensions that have been specifically formulated to evaluate successful trainee behaviour during the SMTA. Existing or new exercises incorporated into the SMTA could be used to provide trainees with the opportunity of demonstrating the required behaviour to be assessed (Gordon & Kleiman, 1976; Mills, 1976; Tziner, 1984), and a training behaviour rating could be allocated to each trainee for the SMTA. A predictive validity study comparing candidate behaviour during the MMAC to trainee behaviour on the SMTA would then fulfil the requirements of Wernimont and Campbell's (1968) consistency theory.

(b) Utilising an assessment centre for which it was not designed

Throughout the literature, researchers and writers have emphasised the necessity of designing an assessment centre specifically for the aim for which it will be utilised, based on a thorough job analysis of the position concerned (Bray & Grant, 1966; Moses, 1977; Thornton & Byham, 1982; Borman, 1982; Tziner, 1984; Spangenberg et al, 1989). The design of the MMAC is based on a job analysis of a manager's position at the next level after middle management, and the dimensions identified and assessed are *those which are required for job success* at this next level. However, the dimension ratings obtained on the MMAC are subsequently used *to predict training success of middle managers on the SMTA*. The MMAC is thus not designed for the purpose for which it is being utilised, namely, selecting managers for entry to the SMTA.

The above reasoning is confirmed by reported research conducted to investigate the validity of assessment centres using different criteria of success. When the assessment centre was designed specifically to predict training success, significant correlation coefficients of 0,53 (Borman, 1982), 0,60 (Tziner, 1984), 0,38 (Tziner & Dolan, 1982) and 0,47 (Mills, 1976) were reported. When an assessment centre was designed to predict job success and validated using training criteria, insignificant correlation coefficients of -0,04 (Bray & Campbell, 1968), -0,11 (Feltham, 1988), -0,29 [Mays (1972), as quoted by Feltham (1988)], and -0,14 (Pynes & Bernardin, 1989) were reported.

In order to test this second explanation it may be necessary to conduct a job analysis of a trainee's position on the SMT<sub>C</sub> to identify those managerial dimensions which are required for success on this course. These dimensions could then be utilised during the MMAC to determine an OAR representing the probability of success on the course, which could be used to determine entry to the SMT<sub>C</sub>. This would ensure that the MMAC is designed to predict successful trainee behaviour, the purpose for which it was originally intended.

(c) **Restriction of range**

According to Cascio (1987) and Dobson (1988), the size of a validity coefficient is a function of two variables, and a narrowing of the range of either the predictor or the criteria scores will result in an underestimation of the true validity coefficient. In the present study, restriction of range was evident in both the criteria and predictor scores.

***Restriction of range in criteria scores.*** As reported in Table 6.2, the variance in the Theoretical, Practical and Total Grade scores is 2.8, 3.9 and 2.8 respectively. The low validities obtained for the OAR could be due in part to these restricted criteria scores (Schmidt et al, 1976). A possible reason for this small variance in scores could be due to the fact that only candidates who obtained an OAR of at least 60 percent are allowed to attend the SMT<sub>C</sub>, and it can be expected that being selected trainees, they should all do reasonably well on the course. To test this explanation it will be necessary to allow all candidates who have attended the MMAC to attend the

SMTc, irrespective of their performance on the MMAC. This should then produce scores on the SMTc which vary more than when only selected trainees are allowed entry to the SMTc.

*Restriction of range in predictor scores.* The research sample of 77 middle managers for this study did not include the 18 candidates who attended the MMAC but did not obtain a required OAR of at least 60 percent. As these candidates were not granted entry to the SMTc, no training grades were available for them, which excluded them from this validity study. This produced what Cascio (1987) refers to as restriction of range in the predicting measures, and, according to Dobson (1988), could result in validity estimates which are an underestimation of the true population validities. Accordingly, the formula for restriction of range (Thorndike, 1949) was applied, and the corrected validity coefficients obtained for the weighted OAR against the three training criteria are -0,03, 0,32 and 0,22 respectively. While the validity of the OAR against the Theoretical Grade remains low ( $r = -0,02$  decreasing to  $r = -0,03$ ), the validities of the OAR against the Practical Grade ( $r = 0,21$  increasing to  $r = 0,32$ ) and the Total Grade ( $r = 0,14$  increasing to  $r = 0,22$ ) are considerably higher than the uncorrected validities.

A possible explanation for the low validity between the OAR and the Theoretical Grade, even after correcting for restriction of range, is the lack of point-to-point correspondence between the behavioural predictor of the MMAC and the cognitive criterion score of the SMTc (Asher & Sciarrino, 1974; Wernimont & Campbell, 1968). When the content

of a managerial training programme is mainly academically oriented, research has shown that the highest predictive validities were obtained when predicting measures such as cognitive ability tests, which have a point-to-point correspondence with the cognitive training criteria, were used (Borman, 1982; Ghiselli, 1973; Gordon & Kleiman, 1976; Mills, 1976; Siegel & Bergman, 1975). When the content of a managerial training programme is mainly behaviour or practically oriented, research has shown that the highest validities were obtained when predicting measures such as work-sample and situational performance tests, which correspond with the practical training criteria, were used (Asher & Sciarrino, 1974; Robertson & Kandola, 1982; Tziner & Dolan, 1982).

The corrected validity coefficient between the OAR and the Total Course Grade ( $r = 0,22$ ) falls within the ranges reported by Gaugler et al (1987) as shown in Table 7.1. This corrected OAR validity supports the widely held contention that assessment centres have predictive validity (Byham, 1970; Howard, 1974; Huck, 1973; Hunter & Hunter, 1984, Thornton & Byham, 1982; Spangenberg et al, 1989).

#### **7.2.2 Moderators of MMAC validity**

Age was shown to correlate significantly with the three training criteria (See Table 6.4), producing correlation coefficients of  $-0,33$ ,  $-0,23$  and  $-0,33$  ( $p < 0,01$ ,  $0,05$  and  $0,01$  respectively). As these correlations are negative, they indicate that the older a candidate, the more likely he is to perform poorly on the SMTc. These results for the MMAC are to be expected, as the SMTc is an intensive course extending over a period of 11 months, exerting

mental and physical pressure on trainees. Younger trainees, who should have a greater capacity for enduring these pressures, should perform better than their older colleagues. Age can thus be considered to be a significant moderator of MMAC validity. This is similar to the findings of Burroughs et al (1973), who found that age was a significant moderator of assessment centre performance, reporting a correlation coefficient of -0,34 ( $p < 0,01$ ).

#### **7.2.3 Accuracy of MMAC predictor combinations**

The results of the linear and stepwise regression analyses indicate that the most accurate predictor of the Total Training Grade obtained on the SMTG, and thus the best available measure to use to determine entry to this course, is a combination of each assessee's final weighted OAR, his age and his experience. This predictor combination produced a multiple correlation coefficient of 0,41 and explains 17 percent of the training criteria variance ( $p < 0,01$ ). These findings are similar to those of Spangenberg et al (1989) and Tziner and Dolan (1982), who also found an increase in assessment centre prediction accuracy by combining various assessment centre predictors.

#### **7.2.4 Validity of weights used to calculate the OAR**

When a mechanical Average Dimension Rating is used, a correlation of 0,13 is obtained with the Total Grade achieved on the SMTG as the criterion of success. When weights are allocated to the consensus dimension ratings and a weighted OAR is calculated, the correlation obtained is 0,14 with the same criterion (See Table 6.3). The use of

these weights do thus not significantly improve the ability of the MMAC to predict training results achieved on the SMTA.

A possible explanation for this lack of increased validity could be due to the fact that the weights were determined with on-the-job performance, and not training success, as the basis. As these two are not necessarily congruent (Tziner & Dolan, 1982), the weights may be invalid for predicting training results, but valid for predicting job performance. To test this explanation it would be necessary to validate the MMAC using job performance as the criterion of success.

### 7.3 LIMITATIONS OF THE STUDY

#### 7.3.1 Size of research sample used

The main limitation of the present study was the size of the research sample used. According to Schmidt, Hunter and Urry (1976), sample sizes required to produce adequate power in validation studies are substantially larger than between 30 and 50, which the authors claim to be a widespread assumption. As an example they report that the average ( $n = 68$ ) validity study has a 50 percent chance of detecting validity if it exists.

According to the tables published by Schmidt et al (1976), which they claim indicate the true sample sizes required in validity studies to detect validity, the sample of 77 managers used in the present study may be inadequate to detect significant validity between the predictors used in the MMAC and the training results achieved on the SMTA if it exists. The correlation coefficients

obtained in this study could thus be underestimates of the true population coefficients.

A possible strategy to reduce the inhibiting effect of the small sample size used in this study would be to follow the suggestion by Cascio (1987), namely, to continue to obtain data on the MMAC and the SMTc and to repeat this predictive validity study at a later date. If similar results are obtained, greater confidence can be placed in the results reported for this study.

### **7.3.2 Reliability of predictors and criteria**

In a meta-analysis of assessment centre validity conducted by Gaugler et al (1987), statistical artifacts such as unreliability of predictors and criteria and range restriction accounted for 54 percent of observed variance of the total research sample. In the present study, validity coefficients were corrected for restriction of range, but it was not possible to determine the reliability of predictor and criterion scores, due to the unavailability of the required data from the organisation in which this study was conducted. The amount of unexplained MMAC validity variance, due to the unreliability of these measures, is thus unknown.

## **7.4 CONCLUSIONS**

In chapter 1 the general aim of this study was formulated as being *to determine whether the MMAC predicts training results achieved on the SMTc*. Following the completed literature review, a research hypothesis was formulated in chapter 5 which read as follows:

**THE ASSESSMENT CENTRE IS A PREDICTOR OF MANAGERIAL TRAINING RESULTS.** Having conducted the empirical study the following conclusions can be made :

**7.4.1 The assessment centre is a predictor of managerial training results**

The corrected validity for the MMAC OAR shows it to be a valid predictor of the Total Grade achieved on the SMTA. In terms of the present study, the research hypothesis is consequently accepted. This confirms the results reported by Borman (1982), Tziner (1984) and Tziner and Dolan (1982), which showed the assessment centre to be a valid predictor of officer training results.

**7.4.2 A combination of predictors is more valid than individual predictors**

At present, only the OAR is used to determine entry to the SMTA. The results of the regression analyses conducted during this study indicate that the most accurate predictor of training results achieved on the SMTA is a combination of each candidate's OAR, age and previous experience in the organisation, producing a multiple correlation coefficient of 0,41. This is a substantial improvement over the validity of the OAR when used on its own. When training costs are high, as in the case of the SMTA which extends over a period of 11 months, any improvement in prediction is worthwhile (Cascio, 1987), and the organisation in question could save money if a combination of predictors are used to eliminate potentially unsuccessful candidates. This conclusion confirms the findings of Tziner and Dolan (1982) regarding the increase in prediction gained by using a combination of predictors.

**7.4.3 Levels of validity for predictors depend on the nature of the criteria used**

The OAR was shown to be an invalid predictor of theoretical training results ( $r = -0,03$ ) but a valid predictor of practical ( $r = 0,32$ ) and total course grades ( $r = 0,22$ ). These results indicate that the levels of validity for the OAR depend on the nature of the three training criteria used when computing the correlation coefficients. This confirms the statements by Borman (1982), Cascio (1987) and Sackett and Ryan (1985) that the nature of criteria influences the levels of validity obtained.

**7.4.4 The use of weights to calculate the final OAR does not improve the ability of the MMAC to predict training results**

The weights used to calculate the final OAR do not lead to a significant increase in the predictive validity of the dimension ratings. The additional time and energy required to calculate a weighted OAR for each assessee is thus unjustified.

**7.5 RECOMMENDATIONS**

The following recommendations regarding the present study can be made:

**7.5.1 Trainees on the SMTC should be assessed according to behavioural dimensions**

Besides the theoretical and practical grades allocated to trainees during the SMTC, they should be assessed according to specific behavioural dimensions which represent successful trainee be-

haviour, and this rating should be used in addition to the theoretical and practical grades allocated to trainees to determine whether they pass the SMT<sub>C</sub> or not.

#### **7.5.2 Dimensions representing successful training behaviour should be determined**

A job analysis should be conducted of the position of a trainee on the SMT<sub>C</sub> to determine the behavioural dimensions essential for success on the course. These dimensions should be used during the MMAC to determine an OAR representing the probability of success on the SMT<sub>C</sub>, and used to determine entry to this course.

#### **7.5.3 Weights should not be used to calculate an OAR**

To save time and money without causing a decrease in the predictive validity of the MMAC, the Average Dimension Rating, and not a weighted OAR, should be used to determine entry to the SMT<sub>C</sub>.

### **7.6 DIRECTIONS FOR FUTURE RESEARCH**

Out of this study, the following directions for possible future research can be given :

#### **7.6.1 Utility of the assessment centre as a selection device for training programmes**

The assessment centre is a labour intensive selection device requiring a substantial financial investment (Pynes & Bernardin, 1989; Thornton & Byham, 1982), and questions have been raised regarding the relative validity of assessment com-

pared with less expensive procedures (Borman, 1982; Cascio & Silbey, 1979). More research is thus needed which employs alternative predictors within the same sample to obtain answers to the assessment centre utility question. In this regard studies by Gardner and Williams (1973), Mills (1976) and Tziner and Dolan (1982), have shown that the use of cognitive ability tests can out perform assessment centres in predicting future managerial performance.

#### **7.6.2 Validating the MMAC with on-the-job performance**

In order to justify the use of an assessment centre as a selection device to determine entry to a training course, it must be validated using either training or on-the-job performance as the criterion of success (Russel, 1984). In this study the MMAC has been shown to be a valid predictor of training results, but it is based on dimensions identified as important for on-the-job performance at the next level after middle management, not training performance on the SMTA. As these two are not necessarily congruent (Tziner & Dolan, 1982), questions have been raised in the organisation in which this study was conducted as to the MMAC's validity in predicting on-the-job performance. A predictive validity study utilising on-the-job performance at the next level after middle management should thus be conducted to provide answers regarding the validity of the dimensions used by the MMAC.

#### **7.7 SUMMARY**

In this final chapter of the report, the research results reported in the previous chapter were in-

terpreted and discussed. The chapter began with a discussion of the predictive validity of the MMAC, after which the limitations of the study were given. Conclusions reached and recommendations made were then reported, and the chapter was concluded by giving directions for possible future research on assessment centres.

## REFERENCE LIST

- Asher, J.J. & Sciarrino, J.A. (1974). Realistic work sample tests: A review. *Personnel Psychology*, 27, 519-533.
- Augustyn, J.C.D. & Van Wyk, A.J. (1988). Die sestien-persoonlikheidsfaktorvraelys as hulpmiddel by die takseersentrum. *Journal of Industrial Psychology*, 14(2), 25-27.
- Banks, C.G. & Roberson, L. (1985). Performance appraisers as test developers. *Academy of Management Review*, 10, 128-142.
- Bateman, T.S. & Zeithaml, C.P. (1993). *Management: Function and strategy*. Boston, MA: Erwin.
- Beatty, R.W. & Schneier, C.E. (1981). *Personnel administration: An experiential skill-building approach* (2nd ed.). Reading, MA: Addison-Wesley.
- Berg, Z.C., Flowers, J., Cilliers, F.V.N., Kriek, H.J., Viviers, A.M., Troost, A.P. & Wolfaardt, J.B. (1980). *Industrial psychological evaluation*. Pretoria: University of South Africa.
- Blanskby, M. & Iles, P. (1990). Recent developments in assessment centre theory, practice and operation. *Personnel Review*, 19, 33-44.
- Boche, A. (1977). Management concerns about assessment centers. In J.L. Moses & W.C. Byham (Eds.), *Applying the assessment center method* (pp. 243-260). New York: Pergamon.
- Borman, W.C. (1982). Validity of behavioral assessment for predicting military recruiter performance. *Journal of Applied Psychology*, 67, 3-9.

Bray, D.W. (1964). The management progress study. *American Psychologist*, 19, 419-429.

Bray, D.W. & Campbell, R.J. (1968). Selection of salesmen by means of an assessment center. *Journal of Applied Psychology*, 52(1), 36-41.

Bray, D.W. & Grant, D.L. (1966). The assessment center in the measurement of potential for business management. *Psychological Monographs*, 80(17), 1-27.

Burke, M.J. & Day, R.R. (1986). A cumulative study of the effectiveness of managerial training. *Journal of Applied Psychology*, 71(2), 232-245.

Burroughs, W.A., Rollins, J.B. & Hopkins, J.J. (1973). The effects of age, departmental experience, and prior rater experience on performance in assessment center exercises. *Academy of Management Journal*, 16, 335-339.

Byham, W.C. (July/August 1970). Assessment centers for spotting future managers. *Harvard Business Review*, 150-165.

Byham, W.C. & Thornton, G.C. (1970). Assessment centers: A new aid in management selection. *Studies in Personnel Psychology*, 2(2), 21-35.

Camp, R.R., Blanchard, P.N. & Huszczo, G. (1986). *Toward a more organizationally effective training strategy and practice*. Englewood Cliffs, NJ: Prentice-Hall.

Campbell, J.P. (1971). Personnel training and development. *Annual Review of Psychology*, 22, 565-602.

Campbell, J.P., Dunnette, M.D., Lawler, E.E. III & Weick, K.R. (1970). *Managerial behavior, performance and effectiveness*. New York: McGraw-Hill.

- Cascio, W.F. (1987). *Applied psychology in personnel management* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Cascio, W.F. & Phillips, N.F. (1979). Performance testing: A rose among thorns? *Personnel Psychology*, 32, 751-766.
- Cascio, W.F. & Silby, V. (1979). Utility of the assessment center as a selection device. *Journal of Applied Psychology*, 64, 107-118.
- Charoux, E. (1990). Assessment centres: A route to black advancement. *Human Resource Management*, 8(6), 30-32.
- Cloete, J.J.N. (1986). *Introduction to public administration* (3rd ed.). Pretoria: Van Schaik.
- Crooks, L.A. (1977). The selection and development of assessment center techniques. In J.L. Moses & W.C. Byham (Eds.), *Applying the assessment center method* (pp. 69-87). New York: Pergamon.
- Cuming, M.W. (1980). *The theory and practice of personnel management* (4th ed.). London: Heinemann.
- Dobson, P.M. (1988). The correction of correlation coefficients for restriction of range when restriction results from the truncation of a normally distributed variable. *British Journal of Mathematical and Statistical Psychology*, 41, 227-234.
- Downs, S. (1970). Predicting training potential. *Personnel Management*, 2, 26-28.
- Dreher, G.F. & Sackett, P.R. (1981). Some problems with applying content validity evidence to assessment center procedures. *Academy of Management Review*, 6(4), 551-560.

Drucker, P.F. (1968). *The practice of management*. London: Pan Books.

Dunnette, M.D. (1976). Aptitudes, abilities, and skills. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 473-520). Chicago: Rand McNally.

Feltham, R (1988). Validity of a police assessment centre: A 1-19-year follow-up. *Journal of Occupational Psychology*, 61(2), 129-144.

Flanagan, J.C. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327-358.

Flowers, J. (1981). *Research methodology*. Pretoria: University of South Africa.

Gardner, K.E. & Williams, A.P.O. (1973). A twenty-five year follow-up of an extended interview selection procedure in the Royal Navy. *Occupational Psychology*, 47, 1-13.

Gatewood, R.D. & Feild, H.S. (1990). *Human resource selection*. (2nd ed.). Orlando, FL: Dryden Press.

Gaugler, B.B. & Thornton, G.C. (1989). Number of assessment center dimensions as a determinant of assessor accuracy. *Journal of Applied Psychology*, 74(4), 611-618.

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

Ghiselli, E.E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26, 461-478.

Glueck, W.F. (1982). *Personnel: A diagnostic approach* (3rd ed.). Texas: Business Publications.

Goldstein, I. (1980). Training in work organizations. *Annual Review of Psychology, 31*, 229-272.

Gordon, M.E. and Cohen, S.I. (1973). Training behaviour as a predictor of trainability. *Personnel Psychology, 26*, 261-272.

Gordon, M.E. & Kleiman, L.S. (1976). The prediction of trainability using a work sample test and an aptitude test: A direct comparison. *Personnel Psychology, 29*, 243-253.

Greenwood, J.M. & McNamara, W.J. (1967). Interrater reliability in situational tests. *Journal of Applied Psychology, 51*, 101-106.

Hall, J. & Donnell, S.M. (1990). Bureaucratic management: The quiet crisis in government. In J. Hall (Ed.), *Models for management: The structure of competence*. (2nd ed.). Woodlands, TX: Woodstead.

Hanekom, S.X. (1987). *Public policy*. Johannesburg: Macmillan.

Helmreich, R., Bakeman, R. & Radloff, R. (1973). The life history questionnaire as a predictor of performance in navy diver training. *Journal of Applied Psychology, 57*, 148-153.

Hersey, P. & Blanchard, K.H. (1993). *Management of organizational behavior*. (6th ed.). Ithaca, NY: Prentice-Hall.

Hinrichs, J.R. (1976). Training techniques. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 829-860). Chicago: Rand McNally.

Hinricks, J.R. (1978). An eight-year follow-up of a management assessment center. *Journal of Applied Psychology*, 63(5), 596-601.

Howard, A. (1974). An assessment of assessment centers. *Academy of Management Journal*, 17(1), 115-134.

Huck, J.R. (1973). Assessment centers: A review of the external and internal validities. *Personnel Psychology*, 26, 191-212.

Human, P. and Human, L. (1989). Towards a better understanding of management development in South Africa. *South African Journal of Business Management*, 20(1), 13-20.

Hunter, J.E. and Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96(1), 72-98.

Jaffee, C.L. (1971). *Effective management selection: The analysis of behaviour by simulation techniques*. Reading, MA: Addison-Wesley.

Jones, A. & Harrison, E. (1982). Prediction of performance in initial officer training using reference reports. *Journal of Occupational Psychology*, 55, 35-42.

Jones, A., Herriot, P., Long, B. & Drakeley, R. (1991). Attempting to improve the validity of a well-established assessment centre. *Journal of Occupational Psychology*, 64, 1-21.

Kane, J.S. & Lawler, E.E. (1978). Methods of peer assessment. *Psychological Bulletin*, 85, 555-586.

Katz, D. & Kahn, R.L. (1978). *The social psychology of organizations* (2nd ed.). New York: Wiley.

- Keil, E.C. (1981). *Assessment centers: A guide for human resource management*. Reading, MA: Addison-Wesley.
- Kerlinger, F.N. (1973). *Foundations of behavioral research*. New York: Rinehart and Winston.
- Kirkpatrick, D.L. (1979). Techniques for evaluating training programs. *Training and Development Journal*, 33(6), 78-92.
- Klimoski, R.J. & Brickner, M. (1987). Why do assessment centers work? The puzzle of assessment center validity. *Personnel Psychology*, 40, 243-260.
- Klimoski, R.J. & Strickland, W.J. (1977). Assessment centers: Valid or merely prescient. *Personnel Psychology*, 30, 353-363.
- Korman, A.K. (1968). The prediction of managerial performance: A review. *Personnel Psychology*, 21, 295-322.
- Kroon, J. (1990). Management and managers. In J. Kroon (Ed.), *General management* (pp. 3-24). Pretoria: Haum.
- McCormick, E.J. & Ilgen, D.R. (1984). *Industrial psychology* (7th ed.). New York: Englewood Cliffs.
- McEvoy, G.M. & Beatty, R.W. (1989). Assessment centers and subordinate appraisals of managers: A seven-year examination of predictive validity. *Personnel Psychology*, 42, 37-52.
- McGehee, W. and Thayer, P.W. (1961). *Training in business and industry*. New York: Wiley.
- Mills, R.B. (1976). Simulated stress in police recruit selection. *Journal of Police Science and Administration*, 4(2) 179-186.

- Mitchel, J.O. (1975). Assessment center validity: A longitudinal study. *Journal of Applied Psychology*, 60, 573-579.
- Moses, J.L. (1977). The assessment center method. In J.L. Moses & W.C. Byham (Eds.), *Applying the assessment center method*. New York: Pergamon.
- Moses J.L. (1987). Assessment centers. In R.L. Craig (Ed.), *Training and development handbook* (3rd ed.). New York: McGraw-Hill.
- Moses, J.L. & Boehm, V.R. (1975). Relationship of assessment center performance to management progress of women. *Journal of Applied Psychology*, 60, 527-529.
- Nagle, B.F. (1953). Criterion development. *Personnel Psychology*, 6, 271-288.
- Noe, R.A. & Schmitt, N. (1986). The influence of trainee attitudes on training effectiveness: Test of a model. *Personnel Psychology*, 39, 497-523.
- Nunns, C. & Kruger, D. (1986). Personnel recruitment and selection. In J. Barling, C. Fullagar & S. Bluen (Eds.), *Behaviour in organisations* (pp. 259-801). Johannesburg: McGraw-Hill.
- Pagonis, W.G. & Cruikshank, J.L. (1992). *Moving mountains: Lessons in leadership and logistics from the gulf war*. Boston, MA: Harvard Business School.
- Peters, T. and Waterman, R. (1982). *In search of excellence*. New York: Harper and Row.
- Pynes, J.E. & Bernardin, H.J. (1989). Predictive validity of an entry-level police officer assessment center. *Journal of Applied Psychology*, 74 (5), 831-833.

Reilly, R.R. & Chao, G.T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 35, 1-62.

Ritchie, R.J. & Moses, J.L. (1983). Assessment center correlates of women's advancement into middle management: A 7-year longitudinal analysis. *Journal of Applied Psychology*, 68(2), 227-231.

Robertson, I.T. & Downs, S. (1979). Learning and the prediction of performance: Development of trainability testing in the United Kingdom. *Journal of Applied Psychology*, 64(1), 42-50.

Robertson, I.T. & Downs, S. (1989). Work-sample tests of trainability: A meta-analysis. *Journal of Applied Psychology*, 74(3), 402-410.

Robertson, I.T. & Kandola, R.S. (1982). Work sample tests: Validity, adverse impact and applicant reaction. *Journal of Occupational Psychology*, 55, 171-183.

Robertson, I., Gratton, L. & Sharpley, D. (1987). The psychometric properties and design of managerial assessment centres: Dimensions into exercises won't go. *Journal of Occupational Psychology*, 60, 187-195.

Russel, J.S. (1984). A review of fair employment cases in the field of training. *Personnel Psychology*, 37, 261-276.

Ryman, D.H. & Biersner, R.J. (1975). Attitudes predictive of diving training success. *Personnel Psychology*, 28, 181-188.

Saari, L.M., Johnson, T.R., McLaughlin, S.D. & Zimmerle, D.M. (1988). A survey of management training and education practices in U.S. companies. *Personnel Psychology*, 41, 731-743.

- Sackett, P.R. (1987). Assessment centers and content validity: Some neglected issues. *Personnel Psychology*, 40(1), 13-25.
- Sackett, P.R. & Ryan, A.M. (1985). A review of recent assessment center research. *Journal of Management Development*, 4(4), 13-27.
- Sakinofsky, I.G. & Raubenheimer, I. van W. (1982). The validity of the assessment centre method as a predictor of managerial potential. *Perspectives in Industrial Psychology*, 8(2), 23-41.
- Saunders, D.R. (1956). Moderator variables in prediction. *Educational and Psychological Measurement*, 16, 209-222.
- Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 66, 178-200.
- Schein, E.H. (1980). *Organizational psychology* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Schmidt, F.L. & Kaplan, L.B. (1971). Composite versus multiple criteria: A review and resolution of the controversy. *Personnel Psychology*, 24, 419-434.
- Schmitt, F.L., Hunter, J.E. & Urry, V.W. (1976). Statistical power in criterion-related validity studies. *Journal of Applied Psychology*, 61, 473-485.
- Schmitt, N., Gooding, R.Z., Noe, R.A. & Kirsh, M. (1984). Meta-analysis of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Siegel, A.I. & Bergman, B.A. (1975). A job learning approach to performance prediction. *Personnel Psychology*, 28, 325-339.

Society for Industrial Psychology (1992). *Guidelines for the validation and use of personnel selection procedures.* Pretoria: Society for Industrial Psychology.

Spangenberg, H.H., Esterhuyse, J.J., Visser, J.H., Briedenhann, J.E. & Calitz, C.J. (1989). Validation of an assessment centre against BARS: An experience with performance related criteria. *Journal of Industrial Psychology*, 15(2), 1-10.

SPSS User's Guide (1983). *A complete guide to the statistical package for the social sciences.* New York: McGraw-Hill.

Stewart, A. & Stewart, V. (1976). *Tomorrow's men today.* London: Western Printing Services.

Stroebel, L. & Raubenheimer, I. van W. (1983). Die verband tussen prestasiemotivering en prestasie aan 'n bestuursbeoordelingsentrum. *Perspectives in Industrial Psychology*, 9(3), 1-10.

Task Force on Assessment Center Guidelines (1989). *Guidelines and Ethical Considerations for Assessment center operations.* Pittsburgh, PA: Human Sciences Press.

Taylor, T.R. (1982). Trainability testing: Pros and cons. *South African Journal of Psychology*, 12(2), 51-56.

Taylor, E.K. & Tajen, C. (1948). Selection for training: Tabulating equipment operators. *Personnel Psychology*, 1, 341-348.

Thorndike, R.L. (1949). *Personnel selection: Tests and measurement techniques.* New York: Wiley.

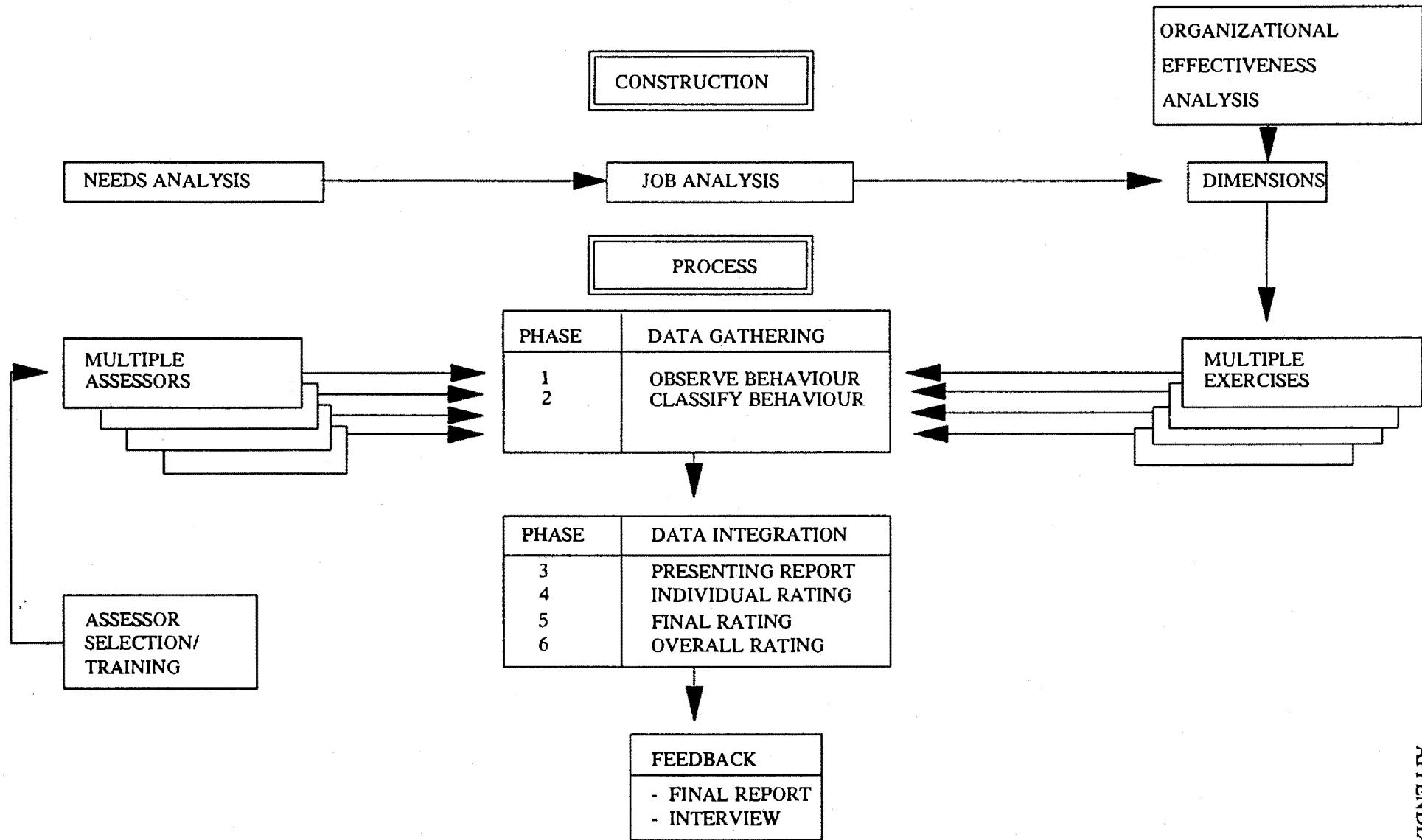
- Thornton, G.C. & Byham, W.C. (1982). *Assessment centers and managerial performance*. San Diegeo, CA: Academic Press.
- Turnage, J.J. & Muchinsky, P.M. (1984). A comparison of the predictive validity of assessment center evaluations versus traditional measures in forecasting supervisory job performance: Interpretive implications of criterion distortion for the assessment paradigm. *Journal of Applied Psychology*, 69(4), 595-602.
- Tziner, A. (1984). Prediction of peer rating in a military assessment center. *Canadian Journal of Administrative Sciences*, 1(1), 146-160.
- Tziner, A. & Dolan, S. (1982). Validity of an assessment center for identifying future female officers in the military. *Journal of Applied Psychology*, 67(6), 728-736.
- Vernon, P.E. (1950). The validation of Civil Service Selection Board procedures. *Occupational Psychology*, 24, 75-95.
- Walker, J.W. (1992). *Human resource strategy*. New York: McGraw-Hill.
- Weislogel, R.L. (1954). Development of situational tests for military personnel. *Personnel Psychology*, 7, 492-497.
- Weiss, D.J. (1976). Multivariate procedures. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 327-362). Chicago: Rand McNally.
- Wernimont, P.F. & Campbell, J.P. (1968). Signs, samples and criteria. *Journal of Applied Psychology*, 52, 372-376.

Wexley, K.N. & Latham, G.P. (1981). *Developing and training human resources in organisations*. Glenview, IL: Scott-Foreman.

Wollowick, H.B. & McNamara, W.J. (1969). Relationship of the components of an assessment center to management success. *Journal of Applied Psychology*, 53, 348-352.

Woodruffe, C. (1990). *Assessment centres: Identifying and developing competence*. London: Short Run.

## STEPS FOLLOWED IN THE DEVELOPMENT OF THE MMAC



## EXAMPLE OF AN EXERCISE - DIMENSION MATRIX

	IN-TRAY	ANALYSIS	GROUP DISCUSSION		INTERVIEW	OBSERVER			CONSENSUS	WEIGHT	MARK
			N-COMP	COMP		1	2	3			
Analytical Ability	6	4			8				6	11	66
Judgment	6	4			7				6	6	36
Decisiveness	2				8				6	5	30
Plan and Organising	6	5			8				6	12	72
Delegating	10								10	9	90
Control	5				7				6	8	48
Group Leadership			8	7					7	10	70
Individual Leadership	8				7				7	7	49
Tenacity			7	8	8				8	3	24
Negotiation Skills			7	7	6				7	4	28
Oral Communication			8	8	7				8	5	40
Written Communication	7	5							6	5	30
										850	583
									%		68,6