

**THE PREDICTIVE VALIDITY OF THE SELECTION BATTERY FOR
TRAINEE PILOTS IN THE SOUTH AFRICAN AIR FORCE**

by

ADEN-PAUL FLOTMAN

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF ADMINISTRATION

in the subject

INDUSTRIAL PSYCHOLOGY

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF MARIÉ DE BEER

JUNE 2002

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to the following people who made this study possible:

- **Our Heavenly Father**, for having provided me with the capacity to complete this study.
- My supervisor **Prof Marié de Beer**, for her skilful guidance, patience and insight.
- **Colleagues** at MPI, for their insightful contributions.
- **Sumari Bolton, Kenny Makgati and Jannie Hartzberg**, for the statistical processing of the data.
- The **South African Air Force (SAAF)**, for the opportunity to undertake this study and for making the data available.
- **Rosanne** my wife, for her inspiration and support in search of a better future.
- My son **Landreth** and daughter **Shanaaz**, who unselfishly shared their father with the latest, albeit temporary addition to the family – this “*skripsie*”.

Student number: **3300-462-5**

I declare that **The predictive validity of the selection battery for trainee pilots in the South African Air Force** is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

.....

(MR A.P. FLOTMAN)

.....

DATE

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
LIST OF TABLES	ix
LIST OF FIGURES	x
SUMMARY	xi
CHAPTER 1: REVIEW OF THE RESEARCH	1
1.1 BACKGROUND AND MOTIVATION FOR THE RESEARCH	1
1.2 PROBLEM STATEMENT	4
1.3 RESEARCH QUESTIONS	4
1.4 AIMS OF THE RESEARCH	5
1.4.1 General aim	5
1.4.2 Specific aims	5
1.5 THE PARADIGM PERSPECTIVE	6
1.6 RESEARCH DESIGN	8
1.7 RESEARCH METHOD	10
1.8 CHAPTER DIVISION	12
1.9 CONCLUSION	12
CHAPTER 2: THE NATURE AND VALUE OF SCIENTIFIC SELECTION	13
2.1 SCIENTIFIC SELECTION	13
2.1.1 Definition of selection	14
2.1.2 Scientific selection	15

2.1.3 The necessity of scientific selection	16
2.1.4 Selection systems and techniques	18
2.1.5 The utility of psychological testing	21
2.2 STEPS IN DEVELOPING A SCIENTIFIC PERSONNEL SELECTION BATTERY	22
2.2.1 Job analysis	24
2.2.2 Setting of worker requirements	25
2.2.3 Development or selection of predictors	25
2.2.4 Assessment of candidates	26
2.2.5 Hiring of an unselected group of candidates	26
2.2.6 Assessment of candidates on actual job performance	26
2.2.7 Correlating scores of all the candidates on the predictors with indices of performance	27
2.2.8 Hiring from among additional applicants only those who obtain certain minimum scores on the predictors	27
2.3 THE VALIDITY OF PSYCHOLOGICAL TESTS	27
2.3.1 Definition	28
2.3.2 The measurement of test validity	28
2.3.2.1 <i>Content validity</i>	28
2.3.2.2 <i>Construct validity</i>	29
2.3.2.3 <i>Criterion-related validity</i>	29
2.3.3 The purpose of validation	30
2.3.4 The evaluation of a validity coefficient	30
2.3.5 Validation in the context of labour legislation	31
2.3.5.1 <i>The Labour Relations Act of 1995</i>	31
2.3.5.2 <i>Legislation on valid recruitment and selection procedures</i>	32
2.3.5.3 <i>Guidelines for assessment practices</i>	33

2.4 THE PROCEDURE OF VALIDATING A SELECTION BATTERY	34
2.4.1 Job analysis	34
2.4.2 Development of criterion measures of job performance	34
2.4.3 Selection of predictors	35
2.4.4 Composition of study sample	35
2.4.5 Statistical analysis	35
2.4.6 Implementation of validity study results	36
2.5 THE PSYCHOLOGICAL PROFILE OF A PILOT IN THE SAAF	36
2.6 CONCLUSION	39
CHAPTER 3: EMPIRICAL STUDY	40
3.1 POPULATION AND SAMPLE	40
3.2 MEASURING INSTRUMENTS	42
3.2.1 The Blox test	43
3.2.1.1 <i>The aim of the test</i>	43
3.2.1.2 <i>Description of the test</i>	43
3.2.1.3 <i>Administration of the test</i>	44
3.2.1.4 <i>Validity and reliability of the test</i>	44
3.2.2 The Advanced Progressive Matrices test	44
3.2.2.1 <i>The aim of the test</i>	44
3.2.2.2 <i>Description of the test</i>	44
3.2.2.3 <i>Administration of the test</i>	45
3.2.2.4 <i>Validity and reliability of the test</i>	45
3.2.3 The Vienna Determination test	46
3.2.3.1 <i>The aim of the test</i>	46
3.2.3.2 <i>Description of the test</i>	46
3.2.3.3 <i>Administration of the test</i>	47

3.2.3.4	<i>Validity and reliability of the test</i>	48
3.3	DATA COLLECTION PROCEDURE	49
3.3.1	Data collection in respect of independent variables	49
3.3.2	Data collection in respect of dependent variable	49
3.4	DATA PROCESSING	50
3.4.1	Correlation coefficient	50
3.4.2	Multiple regression	53
3.5	RESEARCH HYPOTHESES	54
3.6	CONCLUSION	55
CHAPTER 4:	RESULTS	57
4.1	RESULTS OF INDEPENDENT VARIABLE CORRELATIONS	57
4.2	RESULTS OF THE TOTAL RELATIONSHIP OF INDEPENDENT VARIABLES AND DEPENDENT VARIABLES	61
4.3	RESULTS OF INDIVIDUAL RELATIONSHIPS OF INDEPENDENT VARIABLES AND DEPENDENT VARIABLES	63
4.4	INTEGRATION OF RESULTS	65
4.5	CONCLUSION	69
CHAPTER 5:	CONCLUSION, LIMITATIONS AND RECOMMENDATIONS	70

5.1 GENERAL CONCLUSIONS	70
5.2 LIMITATIONS OF THE STUDY	71
5.3 RECOMMENDATIONS FOR FUTURE RESEARCH	72
5.4 CONCLUSION	74
REFERENCES	75

LIST OF TABLES

TABLE	PAGE
2.1 Psychological profile of pilots in the SAAF as measured by selection battery	37
3.1 Distribution of sample by years of training	41
3.2 Composition of the sample according to gender	41
3.3 Composition of the sample according to race groups	41
3.4 Composition of the sample according to home language	42
4.1 Interdependent correlations between independent variables	58
4.2 Pearson correlations between independent and dependent variables	59
4.3 Predictive values of independent variables with dependent variable	61
4.4 Individual beta weights	62
4.5 Correlations between independent variables and dependent variable	64

LIST OF FIGURES

	Page
2.1 Development of a selection battery	23



658.31120968 FLOT



0001814388

SUMMARY
**THE PREDICTIVE VALIDITY OF THE SELECTION BATTERY FOR TRAINEE
PILOTS IN THE SOUTH AFRICAN AIR FORCE**

By

ADEN-PAUL FLOTMAN

SUPERVISOR : Prof M. de Beer
DEPARTMENT : Industrial Psychology
DEGREE : MA (Industrial Psychology)

The sample comprised 92 candidates who have completed the Ground School Phase of the trainee pilot training programme. The independent variables are Raven's Advanced Progressive Matrices Test, the Blox Test and the Vienna Determination Test. The dependent variable is the candidates' results after the Ground School Phase of their training. The results indicated that only the Advanced Ravens test and the Vienna Determination test (Phase 3) correlate positively with the Ground School Phase results of the candidates.

The current battery emphasises cognitive abilities and psycho-motor functioning and does not cover personality traits of prospective candidates. It is recommended that new instruments are included (to cover the measurement of personality traits), that the sample size should be increased by promoting pilot training among previously disadvantaged students and that the current pilot profile should be updated as a matter of urgency.

Key terms:

Validation; psychometric testing; predictive validity; pilot profile; labour legislation; job analysis; job description; scientific selection

CHAPTER 1

REVIEW OF THE RESEARCH

The purpose of this chapter is to provide the background and motivation for the research. It also presents the problem statement, aims, paradigmatic perspective, research design, research method and chapter division.

1.1 BACKGROUND AND MOTIVATION FOR THE RESEARCH

The South African Air Force (SAAF) is primarily responsible for the protection of South African air space. In times of relative peace it provides humanitarian aid during emergencies. However, the SAAF can only meet these objectives to the extent that it has access to the required human, technical and financial resources. The SAAF has been battling to recruit new pilots and to retain existing ones over the last few years. Pilot selection is therefore, an area that needs to be addressed as a matter of urgency. The situation has been exacerbated by the introduction of new objectives for pilot selection in the post 1994 era.

Heunis (1998) maintains that the most important issues to be addressed are the following:

- a. The group currently applying for pupil pilot selection and being selected does not reflect the distribution in the South African population with regards to race, gender, etc.
- b. It would be difficult to address the above issues immediately without compromising selection, safety and training standards that are in accordance with international requirements.
- c. The number of trained psychologists in the South African Military Health Service (SAMHS) with the necessary experience to conduct a selection process (and especially those with an aviation background) is decreasing. Replacing the lost

expertise is problematic, due to a lack of funds to recruit new, experienced psychologists.

- d. Dwindling financial resources have put constraints on training resources. The number of trained flying instructors is decreasing and available flying hours are under threat of reduction.

Pilot selection has to balance the need to select the very best candidates (in the context of limited and constantly decreasing resources) with the growing need for representivity.

One of the main stumbling blocks towards the ideal of representivity of selected candidates, is the fact that a large number of black candidates tend to under perform on the Vienna Determination Test. (This test is discussed at length in chapter 3). A rudimentary evaluation of the Vienna Determination Test revealed that "the black candidates under perform due to the fact that there are no race-based norms for the test, as well as the fact that this group tends not to have the necessary psycho-motor functional ability where machines are involved, specifically regarding this kind of apparatus" (Makgati, 2000, p.2). The current problem is that there is no statistical (research) evidence to indicate whether the current selection battery has predictive validity.

Aspeling (1990) is of the opinion that flying requirements are constantly changing. This makes the selection procedure more difficult. He further maintains that every flight is unique. Weather conditions are constantly changing, with differences in turbulence, crosswinds, temperature, winter and summer conditions and particularly in operational tasks and conditions. What is also important to note is that the mental state of the pilot is also constantly changing. A pilot's mental state can have a great impact on his/her flying ability, particularly during the first phase of training when complex cognitive material and procedures have to be learnt.

Pilots are supposed to be highly skilled, professional individuals. Due to the nature of

the trainee pilot's work and training, it is critical that candidates be identified who suit the profile of a potential pilot. It is therefore, essential that selection instruments are reliable and valid to meet this objective.

The necessity and obligation always exist for selection batteries to be constantly updated and revised to confirm their validity. Failure to ensure this, may result in the selection procedure being unfair and discriminatory towards some candidates. Failure to attract the right candidates would also have implications for the organisation.

For the SAAF, the improvement of the validity of selection instruments would result in the selection of an effective and efficient "pilot pool" and would also be cost effective for the South African National Defence Force (SANDF) and the SAAF in particular.

Despite the fact that the SANDF falls outside the scope of the Labour Relations Act, No 66 (1995) and the Employment Equity Act, No 56 (1998), the Military Psychological Institute, which is responsible for the pre-selection of trainee pilots, is still bound by its professional and moral integrity to ensure that all the test batteries in use are both legally and scientifically defensible. By engaging in this research, the researcher will contribute towards the literature which already exists.

In the light of the critical shortage of pilots in the SAAF and the acquisition of new aircraft within the next few years, it is imperative that the value of the newly acquired aircraft be enhanced by the selection and training of the most suitable personnel.

If the Military Psychological Institute intends to continue functioning as a leading specialist behavioural sciences institute, it will have to validate its selection batteries on a regular basis. This research project is an attempt to make a contribution in this regard.

The research will also contribute towards the establishment of what is known as a *feedback loop* (Aspeling, 1990). A feedback loop monitors the predictive validity (i.e. the correlation between the scores on the selection instruments and the criterion

scores) of the selection system. According to Aspeling (1990), the predictive validity of all selection systems will decrease over time without intervention. This decrease is normally caused by changes in the applicant pool, changes in the criterion, or compromising of the instruments. The feedback loop will allow the SAAF to determine when the predictive validity of its pilot selection battery is decreasing and to take the most appropriate countermeasures.

1.2 PROBLEM STATEMENT

The current problem the SAAF faces is that there is no statistical evidence to confirm the predictive validity of the current selection battery. It is hoped that this kind of information will help to reduce the high drop out rate of candidates during the Ground School Phase of their training.

1.3 RESEARCH QUESTIONS

Based on what has been presented, the following research questions are formulated:

- What is the value and nature of scientific selection?
- What is involved in the procedure of developing a scientific selection battery?
- What is validity and in particular predictive validity; how does one undertake a statistically based validation study?
- What is the current profile of a pilot in the SAAF?
- What constructs are measured by the current selection battery?
- What are the dependent and independent variables of the study?
- What is the predictive validity of individual tests (which are part of the battery for trainee pilots in the SAAF), as well as the total selection battery, in order to determine training success?
- What are the limitations of the study and what recommendations can be made in terms of selecting the most suitable candidates to be admitted to the trainee pilot training programme?

1.4 AIMS OF THE RESEARCH

The following aims are formulated from the above questions.

1.4.1 General aim

The general aim of the research is to determine whether the existing selection battery for trainee pilots in the SAAF is a valid predictor of training success by correlating the candidates' pre-selection results, with their Ground School Phase results.

1.4.2 Specific aims

In terms of the *literature review* the specific aims are:

- To define and discuss the nature and value of scientific selection
- To discuss the procedure of developing a scientific selection battery
- To define validity and in particular predictive validity and how to undertake a statistically based validation exercise
- To discuss the current profile of a pilot in the SAAF
- To determine the constructs/dimensions that are measured by the current test battery
- To discuss the dependent and independent variables of this study

In terms of the *empirical study* the specific aims are:

- To statistically determine the predictive validity of specific tests (which are part of the battery for trainee pilots in the SAAF), as well as the total selection battery, in order to determine training success
- To discuss the limitations of this study and to formulate recommendations based on the above study in order to improve the current selection battery

1.5 THE PARADIGM PERSPECTIVE

Research is always conducted within the context of a specific paradigm (Mouton & Marais, 1990). It plays a critical role in terms of demarcating the boundaries of the research and to formulate specific points of departure for the research. With reference to the paradigm perspective, the disciplinary relationship of the research, its most applicable paradigms and meta-theoretical assumptions are discussed.

The research is situated within the field of industrial psychology and its field of application. Specific emphasis is placed on personnel psychology and psychometrics.

Industrial psychology is the “scientific study of human behaviour in the production, distribution and consumption of the goods and services of society and it refers to a branch in applied psychology, a term covering organisational, military, economic and personnel psychology” (Reber, 1988, p.352). The tasks of the industrial psychologist include the study of organisations and organisational behaviour, personnel recruitment and selection, human resource management, the study of consumer behaviour, research and psychological testing (Du Toit, 1989).

Personnel psychology is “concerned with all aspects of applied individual differences. Among other things, personnel psychologists determine what human skills and talents are needed for certain jobs, how to assess potential employees, how to grade employee job performance, and how to train workers to improve job performance” (Muchinsky, 1993, p.5).

Plug, Meyer, Louw and Gouws (1986, p.296) refer to *psychometrics* as “the study of aspects of psychological measurement that focus on the development and implementation of mathematical and statistical procedures”. Psychometric tests are “objective standardised measurements of a certain area in human behaviour” (Smit, 1986, p.19).

The literature review is presented within the behaviouristic paradigm. This paradigm is encapsulated in a number of concepts (Chmiel, 2000; Cilliers, 1985; Louw, 1986; Meyer, Moore & Viljoen, 1990) namely:

- *Atomism* (all of reality is made up of elementary components);
- *Materialism* (all phenomena can be reduced to physical matter);
- *Naturalism* (all aspects of human life are natural phenomena and do not differ in any way from other aspects of nature);
- *Positivism* (only observable phenomena can be scientifically studied); and
- *Empiricism* (knowledge is possible only with the aid of sensory perception).

Watson's (the father of behaviourism) viewpoint can be summarised as follows (Meyer et al., 1990):

- the objective observation of behaviour is considered to be the only acceptable method of conducting research;
- behaviour is regarded as consisting of connections between stimuli and responses;
- the duty of psychology is described as the study of stimulus – response connections and the way in which such connections are acquired;
- classical conditioning is regarded as the most important learning method.

For the purposes of this study, the psychometric test could be regarded as the stimulus followed by a response, for example, answering the test questions.

The empirical study is presented from the functionalist paradigm, which is essentially a quantitative approach. The following are basic assumptions of the functionalist paradigm (Meyer et al., 1990; Morgan, 1980, p.608):

- the functionalist perspective is primarily regulative and pragmatic in its basic orientation;
- it is concerned with understanding society in a way that generates useful, empirical knowledge;
- society has a concrete, real existence and a systematic character which is orientated to producing an ordered and regulated state of affairs;
- it encourages an approach to social theory that focuses upon understanding the role of human beings in society; and
- behaviour is always seen as being contextually bound in a real world of concrete and tangible social relationships.

1.6 RESEARCH DESIGN

A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure (McCall, 1990; Selltiz, Jahoda, Deutsch & Cook, 1965).

The quality of a good researcher is that he/she will always attempt to eliminate all those variables that might have an influence on the validity of the results. The research design has a critical role to fulfill in this regard. It helps to enhance the internal and external validity of the research findings (Mouton & Marais, 1990).

The literature review will be presented in a qualitative and the empirical study presented in a quantitative, descriptive way.

In order to ensure internal validity on a contextual level, the researcher will use measuring instruments that are currently used in selection within the SAAF. External validity will be ensured, by using the entire population.

The purpose of the research design is to determine whether the identified independent variables have an impact on the identified dependent variable (Huysamen, 1980). In this research, the independent variables are Raven's Advanced Progressive Matrices Test, the Blox Test and the Vienna Determination Test results. The dependent variable is the candidates' results after the Ground School Phase of their training.

The unit of analysis is the group who successfully completed the psychometric phase of the selection procedure, and who either successfully or unsuccessfully completed the Ground School Phase of their training between 1997 and 1999.

In many validation samples, predictor instruments are administered to groups of employees who have already been employed in a particular job for some time, or to a group of employees who have recently been hired and for whom criterion measures of job performance will soon be available. In both these cases, the chances are that the sample will represent proportionately more persons who possess a "lot" of the characteristic measured by the test than would a sample of those who initially applied for the job. Hence, the range of such a group will be curtailed at the lower end of the score distribution. In other words, the distribution will be negatively skewed (Huysamen, 1980; Muchinsky, 1993; Tabachnick & Fidell, 1996). The effect of this pre-selection will be to lower the validity coefficient from what it would have been, if the original applicants had been used as the validation sample (Anastasi & Urbina, 1997; Cascio, 1991).

In view of this effect, detailed information needs to be supplied about the degree of pre-selection that has occurred. This requires knowledge of the means and range of scores both in the sample in question and in a random sample of job applicants (Reitz, 1981; Saunders, 2000; Van Wyk, 1993).

One of the effects of pre-selection on results through restriction of range is the possible lowering of correlation results. In this study, the optimal psychometric procedure would have been to include *all* candidates in the training. However, there are serious practical limitations to be considered, e.g. safety requirements, finance, etc.

1.7 RESEARCH METHOD

The research was conducted in two phases, each consisting of various steps.

Phase One: Literature review

Phase one entails the literature review, which will be presented in a qualitative and descriptive way (conceptualisation). The steps covered by the literature review are:

- Step 1: A literature review on the nature and value of scientific selection
- Step 2: The procedure of developing a scientific personnel selection battery
- Step 3: The validity of psychological tests
- Step 4: The procedure of validating a selection battery
- Step 5: The psychological profile of a pilot in the SAAF

Phase Two: Empirical research

Phase two entails the empirical study that will be presented in the form of an explanatory study. The steps covered by the empirical study are as follows:

- Step 1: Description of the population and sample

The population comprises all candidates who have been selected for pilot training in the South African Air Force between 1997 and 1999. The sample comprises 92

candidates who have completed the Ground School Phase (18 months) of the training programme.

Step 2: Description of instruments and criterion

Description of the psychometric instruments and the composition of marks during the Ground School Phase.

Step 3: Data collection

The psychometric results and Ground School Phase results of the candidates between 1997 and 1999 were used for the study. The candidates were tested at the Military Psychological Institute (MPI) in Pretoria, where the psychometric results were entered into a database.

Step 4: Data analysis

Data was processed, by using the Statistical Package for Social Science (SPSS) software programme (SPSS Base 9.0, 1999). Data was analysed by means of correlations and regression analysis.

Step 5: Formulation of the empirical hypothesis

Hypotheses are formulated to cover the objectives of the research.

Step 6: Reporting and interpretation of the results

Step 7: Integration of results

Step 8: Formulation of the conclusions of the research

The conclusion is presented in such a way that the objectives of the study are addressed.

Step 9: Limitations of the research

The limitations of the literature review and the empirical study will be discussed.

Step 10: Recommendations

Recommendations will be presented to address the identified problem statements.

1.8 CHAPTER DIVISION

The chapters of the research will be presented in the following manner:

Chapter 2: Literature review

Chapter 3: Empirical study

Chapter 4: Results

Chapter 5: Conclusions, limitations and recommendations

1.9 CONCLUSION

In this chapter the following was discussed: background to and motivation of the study, problem statement, aims, paradigmatic perspective, research design, research method and chapter division.

Chapter two deals with phase one of the study, namely the literature review concerning the nature and value of scientific selection.

CHAPTER 2

THE NATURE AND VALUE OF SCIENTIFIC SELECTION

The purpose of this chapter is to undertake a theoretical study in order to reach a more comprehensive perspective of existing literature in the use of psychometric tests as a scientific selection tool. Reference is made to a selection system, various selection techniques and the value of scientific selection in the field of industrial psychology.

2.1 SCIENTIFIC SELECTION

The explosion in information technology and increasing automation, have had two direct implications on selection. Firstly, it has increased the need for valid tests of complex operational aptitude and secondly, has made the cost of placing the wrong person in charge, greater than ever. These complexities, have given birth to what is generally known as *situational awareness* (Aspeling, 1990).

In the aviation context, situational awareness becomes critical as high volumes of information compels the operator viz. a pilot, to be more attentive, and automation makes it more important to know what is happening at all times and subsequently, to make critical interventions, whenever these are necessary (Heunis, 1998). The pilot is becoming more and more a systems manager. He/she is thrust into action, whenever the system malfunctions and a corrective intervention has to be made. The ideal is to have an individual with the required skills and knowledge including the ability to apply these skills decisively in the confusion of a complex, stressful event. In the context of pilot selection, operational aptitude testing is imperative (Aspeling, 1990).

All organisations need human resources in an attempt to reach their objectives. The organisation has to engage in a process of matching the skills, knowledge, interests and general capabilities of an individual with the specific requirements or competencies of a specific job, as well as the needs and culture of the organisation. The assumption is drawn that the closer the match between these individual characteristics and job

requirements, the more successful and ultimately effective the organisation will be in achieving its objectives (Cascio, 1991).

2.1.1 Definition of selection

Brown, (cited in Smit, 1990, p.12) defines selection as "...a systematic procedure for measuring a sample of behaviour." Cascio (1985, p.19) simply defines selection as "...the central phase in the process of matching individual and job."

For the purposes of this research, the primary aim of selection is therefore, to find the most suitable candidate for a specific job. It is to everybody's advantage (the individual, colleagues, the organisation and society) that the most suitable candidate for a specific job is selected. It is reasonable to assume that a mismatch between the individual and the job would result in the individual being unhappy and ineffective in his/her job, thereby making it difficult for the organisation to reach its objectives (Anastasi & Urbina, 1997; Chmiel, 2000; Cronbach, 1994; Meiring, 1995).

Du Toit (1989) maintains that the selection procedure is based on two assumptions, namely,

- There are individual differences between employees. This has two implications. Firstly, individuals have different abilities with regard to the execution of the same tasks. Secondly, an individual has different abilities with regard to the execution of different tasks.
- Posts differ with regard to content. The employee who occupies the post has to meet specific expectations.

These two assumptions are important to consider in the procedure of finding the best possible person for a specific job. Hence, the importance and value of selection.

2.1.2 Scientific selection

For any selection procedure to be scientific, it has to adhere to the following requirements (Herholdt, 1977; Howell, 1995):

- The use of reliable predictive measures.
- It should be inductively appropriate.
- Its predictive accuracy must be known.

Smit (1990) adds to these criteria by indicating that reliability and accuracy are central elements in the selection procedure. Psychometric tests have to meet the following specific core requirements in order to be meaningful, legally defensible and scientific:

- *Objective*
A test is objective if a candidate is not aware of how his/her test behaviour could influence the interpretation of the test and ultimately the test results (Smit, 1990).
- *Standardised*
Uniformity of administration, scoring and interpretation procedures (Anastasi, 1988; Anastasi & Urbina, 1997; Cronbach, 1994).
- *Construct definition*
It is essential that the designer of the test accurately and clearly defines the construct(s) which the test is supposed to measure (Jordaan & Jordaan, 1986; Lundin, 1996; Morgan, 1980; Smit, 1990).
- *Normal distribution*
The test performance should form a normal distribution if the standardisation sample is sufficiently representative (Smit, 1990; Walton, 1987).

- *Reliability*

If a candidate performs consistently in a test on different occasions, the test is said to be reliable (Jones, 1995; Vinchur, 1993; Wheeler, 1993).

- *Validity*

A test is said to be valid if it measures the abilities, skills, and worker characteristics related to the job in question (Borack, 1994; Gerber, Nel & Van Dyk, 1998; Smit, 1990).

- *Utility*

All the practical considerations for the use of a test are indicative of the usefulness or the utility value of a test (Ree, 1994; Smit, 1990).

All the above-mentioned criteria have particular relevance in South Africa where the population consists of a variety of groupings i.e. ethnic, political, economic and social. Legislation is in place e.g. the Labour Relations Act 66 (1995) and the Employment Equity Act 56 (1998) to ensure that tests and test administrators comply with a number of requirements in an attempt to protect these different groupings from unfair discrimination (Bendix, 1996).

2.1.3 The necessity of scientific selection

It is important to make a clear distinction between a screening system and a selection system. A screening system is normally used to eliminate applicants who do not meet the minimum requirements specified and normally precedes the selection system (Gerber et al., 1998). An applicant is expected to successfully pass the screening process before participating in the selection process. Applicants who do not meet the minimum requirements are therefore, identified and eliminated from further consideration (Cascio, 1991; Louw, 1991; Smit, 1990).

The industrial psychologist is primarily concerned with the assessment of those individual abilities and traits that have some bearing on an individual's level of job

performance (Reber, 1988). The effective utilisation of human resources would be difficult without a comprehensive knowledge of the abilities and traits of individuals and how they manifest themselves in different work situations. This knowledge is normally gained through some form of psychological assessment.

An organisation's human resources are arguably its most valuable asset. The welfare of the organisation is thus closely related to the quality of its personnel. It is assumed that an employee with a high quality of work life would be an asset to the local community and the country as a whole. In order to select and place personnel, they should be assessed in some way. The more valid and reliable this evaluation process, the better the selection and the greater the benefits to the organisation (Klinvex, 1999).

A thorough understanding of how people differ is essential for the suitable placement of personnel. Only if the abilities and traits of employees have been accurately assessed, can the industrial psychologist make suggestions as to the kinds of structures, styles of leadership, patterns of communication networks, and chains of command that will promote high levels of satisfaction and effective performance throughout the organisation (Jones, 1995; Scheffler, 1991).

The accuracy or failure of an instrument to distinguish between the most suitable candidate will have profound implications for the individual and the organisation.

An employee who has been unsuitably placed would find it difficult not only to develop him/herself personally and professionally, but it would also take an exceptional effort to make a positive contribution to the organisation (Van der Walt, 1997).

The meaningfulness of the person's working experience would be gravely diluted and his/her personal sense of pride to be associated with the organisation, adversely influenced (Hoffman & McPhail, 1998).

Van der Walt (1997) outlines the consequences non-scientific and inaccurate selection could have for the individual. She states that psychological damage could occur in the form of a loss of pride, a loss of self-confidence and feelings of inferiority. Other implications include, dissatisfaction, boredom, frustration, demotivation and restlessness.

The appointment of unsuitable personnel could possibly result in the following: organisations battling to achieve their goals; accidents; financial losses; high employee turnover; high rate of personnel absenteeism; and unsuccessful organisational strategies and programmes (Forster, 2000; Jones, 1995).

Most, if not all competitive organisations, exist and strive to be as productive as possible at the lowest possible cost (Van der Walt, 1997). Employees play a critical role in the organisation's effort to achieve this goal and to perform optimally (Van der Walt, 1997). It is costly to invest in the training of individuals who fail to reach criterion performance levels in training or, worse yet, pass all training tests but then are unable to perform up to standard under operational stress. As so often happens with some trainee pilots or Air Traffic Controllers, the individual may have all of the skills and knowledge normally required but is unable to put them together in the confusion of a complex incident (Aspeling, 1990). The scientific selection of personnel therefore, becomes imperative.

In order to identify suitable personnel, one has to start with a proper job analysis and job description. It is only once the requirements of a specific job have been identified that a selection system can be put in place.

2.1.4 Selection systems and techniques

A selection system consists of five major elements (Cascio, 1991; Klinvex, 1999; Muchinsky, 1993; Smit, 1990).

According to these authors the first is the *selection instruments*, for example, written tests, simulator evaluations and interviews.

The second element is the *procedures* used to administer the instruments. These procedures include such things as the order in which the instruments are administered, the time allotted to complete the instrument, the number of people on the interview panel, the media used to present the instruments, and the selection model.

The third element is the *job performance criterion*. The criterion represents the behaviour the selection system is designed to predict. All selection systems must have a criterion that is explicit, comprehensive, and quantifiable. It is the most important element, and often the most neglected.

The fourth element is the *statistical techniques*. These techniques are used to correlate scores of the instruments and the criterion. Statistical techniques also are used to develop cut-off points and maximise the predictive utility of the instruments.

The *feedback loop* is the fifth element of the system. It monitors the predictive validity (the correlation between the scores on the instruments and the scores on the criterion) of the selection system. This feedback loop allows the organisation to determine when the predictive validity is decreasing and to take the appropriate countermeasures. In countries where employment litigation is likely, for instance South Africa, a feedback loop establishes "intent to improve" and provides an additional measure of protection against litigation.

A variety of selection techniques have been developed over the years (Muchinsky, 1993). Some have stood the test of time, while others were discarded along the way. Saunders (2000) distinguishes between four methods through which selection can be done:

- *Application blanks*

Application blanks are generally used in industry. It is used to collect personal information about the candidate including his/her work history.

- *Interviews*

Interviewing is often used for deciding who should be selected initially, who should be placed in different jobs, who needs counselling and training and whose services should be terminated (Hoffman & McPhail, 1998; Van der Walt, 1997). It is an opportunity for personal contact between the candidate and the employer, or his/her representative. However, it could be time-consuming and laborious. The selection interview has a number of primary functions: to determine the candidate's experience, level of training, and its applicability or relevance; to evaluate the candidate's personality characteristics and cognitive functioning.

- *Psychometric testing*

Psychometric tests and inventories have been the most frequently used predictors in I/O Psychology (Muchinsky, 1993). In tests, the answers are either right or wrong, while in an inventory there are no right and wrong answers. Generally, the terms *test* and *psychological testing* are representative of the family of tests and inventories (Anastasi & Urbina, 1997; Robbins, 1998).

Society has tended to imbue psychological tests with mystical powers (Muchinsky, 1993). However, these tests (intelligence, aptitude, ability, personality, etc.) assist the practitioner in making better decisions than could be made without them.

Testing has its place in the repertoire of instruments at the practitioner's disposal. It is common knowledge that some tests are useful in predicting job success and others are not. Hence, the importance of validating the current trainee pilot selection battery.

Saunders (2000) maintains that psychometric testing assists the practitioner in arriving at a more global impression of the candidate. This is due to the fact that more information is generated through psychometric testing than any other method. Despite legislation regulating the use of psychometric material, psychometry remains one of the most popular selection techniques.

- *References*

References can be assessed telephonically, in written form, or personally. A major advantage of this technique is that it is cost effective due to the fact that credibility of a reference can be assessed very quickly. However, one should ensure that the candidate's privacy is not violated.

2.1.5 The utility of psychological testing

The utility of psychological tests can be presented as follows (Abrahams & Alf, 1993; Owen & Taljaard, 1996, p.12):

- To lend objectivity to our observations;
- to elicit behaviour under relatively controlled circumstances;
- to sample the behaviour people are capable of;
- to measure the progress made with regard to set objectives or standards;
- to give insight into aspects of human beings that are not directly observable;
- to trace characteristics and components of behaviour;
- to predict future behaviour; and
- to provide information for feedback and decision-making.

The above is an indication that psychological tests have potential utility value, depending on the purpose and context in which these instruments are used. Selection systems and techniques should therefore, be selected with great care and sensitivity.

2.2 STEPS IN DEVELOPING A SCIENTIFIC PERSONNEL SELECTION BATTERY

A number of steps should be adhered to in developing a scientific selection battery (Lewis, 1990). In order to comply with current labour legislation, it is important to have a clear understanding of legal requirements. Figure 2.1 presents the steps involved in developing a selection battery.

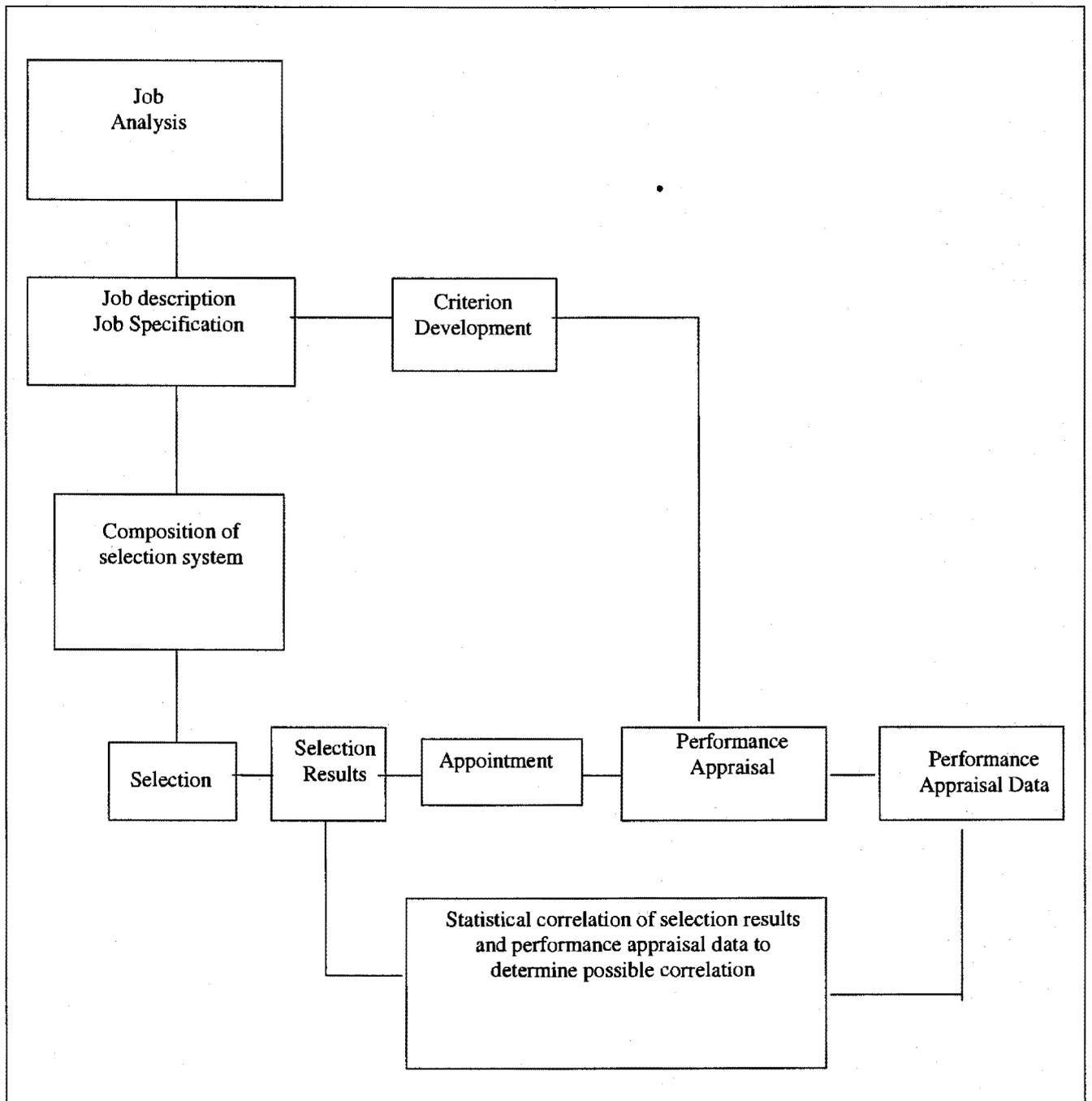


FIGURE 2.1 Development of a selection battery

(Milkovich & Glueck, 1985)

Lewis (1990) suggests that the following steps be used when developing a selection battery:

2.2.1 Job analysis

On the basis of a thorough job analysis, a comprehensive job description is developed of what employees actually do on the job. The primary aim is to develop an accurate understanding of the true nature of the job one wishes to select for. Thus the essential characteristics for the successful performance in a position are identified. Access to this information reduces the possibility of employees being appointed at random and provides the practitioner with a set of criteria for successful work performance (Lewis, 1990).

Jones (1995, p.154) defines a job analysis as follows:

A job analysis is a set of objectives and systematic techniques which provides detailed information about job design and key success criteria.

Alternatively, it is defined as,

A rigorous, highly sophisticated and systematic procedure in collecting accurate qualitative and quantitative information on exactly what the "job" demands and involves, about the individual at work and the job's relative value within an organisation in terms of job content and objectives (i.e. key tasks, competencies needed, responsibilities, pay, error raters, frequency of acts, time spent, importance, decisions to be made, selections and appointments) (Vinchur, 1993, p.145).

For the purposes of this research a job analysis can therefore be defined as a systematic, scientific technique or procedure, which is used for the collection of accurate information about a specific job for a variety of purposes, for example, selection, appointment, development, etc.

Vinchur (1993) further explicates that one of the first principles of testing is that tests must be selected or designed on the basis of a sound job analysis programme. Since the purpose of testing is to predict future success in the job situation, the starting point of analysis is obviously the job.

Gatewood and Field (1987, p.172) offer some specific comments regarding the importance of an accurate job analysis in designing a scientific selection programme:

What is important is the recognition that it is the job analysis process that impacts on the effectiveness of any human resources selection system. Where job analysis is incomplete, inaccurate, or simply not conducted, a selection system may be nothing more than a game of chance - a game that employer, employee, and job applicants alike may lose.

What has been presented so far seems to indicate that the success of a selection battery to distinguish between candidates, rests to a large extent on the accuracy of the initial job analysis.

2.2.2 Setting of worker requirements

Lewis (1990) further explains that by studying the comprehensive job description it is important to hypothesise what kind of individual would be most likely to carry out the behavioural demands of the job effectively. This could be done by describing the person requirements in terms of fairly well-established psychological dimensions and then to verify these descriptions by either testing job incumbents with carefully developed measures of these dimensions or getting the agreement of qualified judges as to their relative importance.

2.2.3 Development or selection of predictors

On completion of this list of well-defined skills, abilities and traits that are considered to be necessary for job success, the next logical step is to produce measures to

assess individual differences in these identified characteristics. Some of the most common of these measures (predictors) are ability tests, projective personality scales, objective personality scales, biographical questionnaires, interviews and situational tests (Lewis, 1990; Muchinsky, 1993; Saunders, 2000; Smit, 1990).

2.2.4 Assessment of candidates

The candidates must then be measured on the predictors (i.e. the selected test instruments). In other words, scores must be obtained for them on the relevant measures. It is necessary to obtain sufficient numbers of candidates so that adequate variation on the predictors is obtained, otherwise it is not possible to determine whether the selection procedure offers any improvement over chance.

2.2.5 Hiring of an unselected group of candidates

Initially, applicants should be hired without regard to their scores on the predictors. This implies that the predictors developed or selected according to the above steps are not utilised in the initial appointment of candidates. The pre-selection of samples leads to restriction of range, thereby lowering correlation results.

2.2.6 Assessment of candidates on actual job performance

After the candidates have been on the job for a sufficient period of time for consistent differences to emerge, their performances are measured. These measures of performance act as the criteria against which the candidates' predictor scores are compared. These criteria must be reliable and valid and they should accurately reflect differences in actual job performance.

2.2.7 Correlating scores of all the candidates on the predictors with indices of performance

This comparison between the performance of candidates who obtain “high” and “low” scores on the various predictors is necessary in order to determine whether the predictors are capable of predicting subsequent job performance. If the relationship, usually expressed as a validity coefficient, is acceptable according to specified standards, the next step is implemented (McCall, 1990).

2.2.8 Hiring from among additional applicants only those who obtain certain minimum scores on the predictors

Once a positive relationship has been established between the predictors and the criteria of performance, it is possible to improve the accuracy of selection by choosing only candidates whose scores are either the same, or higher than those job incumbents, selected without consideration of their predictor scores, who perform satisfactorily on the job.

To establish this cut-off score, it is necessary to compare the scores on the predictors of those present incumbents who are rated as successful and unsuccessful (Tabachnick & Fidell, 1996).

These steps together constitute the “traditional model” of personnel selection.

2.3 THE VALIDITY OF PSYCHOLOGICAL TESTS

The nature of psychological tests demands that the administrator have an indication of what the test actually measures. The basic notion behind validity is the notion of job relatedness.

2.3.1 Definition

A test is said to be valid if it measures the abilities, skills, and worker characteristics related to the job in question (Anastasi & Urbina, 1997; Smit, 1990). A test is therefore valid if persons who score low on the test tend to perform poorly on the job and persons who score high on the test tend to perform successfully on the job. There is therefore a positive correlation between test performance and job performance.

The procedure of validation is never-ending (Huysamen, 1980). Validity studies provide practitioners with critical information about the relationship between a specific test and other observable job-related behaviour (Kerlinger, 1986). This information enables the practitioner to make inferences about present and future job behaviours (Lewis, 1990).

2.3.2 The measurement of test validity

According to Kline (1999) a validation study is an investigation that compares scores on a test or tests with important, on-the-job performance measures. The study will result in a mathematical index of the strength of the relationship between these two, such as a correlation coefficient or a hit-rate percentage. There are three methods of validating a test for its job relatedness: content, construct, and criterion related validity (Kerlinger, 1986):

2.3.2.1 *Content validity*

A measure is said to have content validity if the sample of items in it is fully representative of the universe of questions that could have been asked, or the representativeness or sampling adequacy of the content, the substance of a measuring instrument (Babbie, 1989; Prinsloo, 1992). It is evaluated by showing how well the content of a test samples the subject matter or kinds of situations that it is intended to assess. For example, if a test is claimed to measure an applicant's

spelling ability, the items must be of such a variety and range that we are entitled to claim that the test gives an indication of the applicant's spelling ability in general, not just his ability to spell the few words that "happen" to be included in the test (Smit, 1990).

2.3.2.2 *Construct validity*

A test is said to have construct validity if it measures the hypothetical psychological/theoretical entity (construct) that lies "behind" the candidate's responses and is claimed to be measured by the test, or simply the extent to which a test measures the theoretical construct it is supposed to measure (Kerlinger, 1986). If this "construct" is embedded in a certain theoretical framework, then establishing the construct validity of a test involves the practitioner in testing the theory underlying it.

2.3.2.3 *Criterion - related validity*

Huysamen (1980) distinguishes between concurrent and predictive criterion-related validity. He defines ***predictive*** validity as the accuracy with which a test or instrument enables you to predict some or other future behaviour or status of individuals. It is therefore an estimate of the future performance of a person on some criterion from scores on the test or predictor. This method involves the comparison of test scores at one point in time (e.g. time of appointment) and criterion scores obtained at a later point in time (e.g. after 18 months). A major advantage of this method is that results are directly useful for the selection of future job applicants.

Despite the advantages of this strategy it involves a substantial time delay before results are available. Since the strategy dictates that all applicants be hired irrespective of their scores, it could have certain detrimental consequences for the organisation. In order to negate some of the above-mentioned limitations, validation studies are often conducted concurrently.

Concurrent validity on the other hand, refers to the accuracy with which a test can identify or diagnose the current status of an individual's behaviour. It involves the calculation of a correlation coefficient, but with no time lag between the obtaining of predictor and criterion scores (Babbie, 1989; Ghiselli, 1984; Huysamen, 1980; Owen & Taljaard, 1996).

2.3.3 The purpose of validation

In Industrial Psychology, the procedure of validation serves the following purpose (Herholdt, 1977):

- to determine the predictive validity of specific instruments;
- to eliminate those instruments with low correlation or which tend to duplicate other instruments;
- to serve as basis in order to allocate weights to specific instruments; and
- to determine cut-off points

2.3.4 The evaluation of a validity coefficient

According to Owen and Taljaard (1996) there are three factors to consider when a validity coefficient is evaluated. Firstly, there is the possible attenuation of the validity coefficient due to a restricted range of test scores for the group of candidates for which the coefficient has been determined. The restriction develops as a result of a prior selection of candidates with an instrument related to the present selection tool and/or criterion. A greater restriction results in an attenuated validity coefficient.

The second factor that should be considered is what is known as the base ratio, i.e. the proportion of persons who comply with the minimum criterion requirements according to a prior selection strategy. The assumption is that the larger the base

ratio, the larger the validity coefficient must be for the new selection strategy to result in a given increment in the proportion of successfully selected candidates.

Another factor for consideration is the selection ratio. According to Muchinsky (1993) the selection ratio is the number of job openings divided by the number of job applicants. The proportion of successful selected candidates (based on the criterion) can be increased by selecting a smaller, but according to the predictor, more promising group of candidates.

2.3.5 Validation in the context of labour legislation

Validity plays an increasingly important role in the context of current labour legislation. Legislation has been advanced to ensure that appropriate assessment methods are selected and administered in compliance with specific standards. The most important legislation include, the Labour Relations Act 66 (1995) and the Employment Equity Act 56 (1998).

2.3.5.1 The Labour Relations Act (LRA) of 1995

a. Purpose of the Labour Relations Act (1995)

The Labour Relations Act 66 (1995) states that the overall purpose of the act is “the advancement of economic development, social justice, labour peace and the democratisation of the workplace “ (p.13). It intends to achieve this aim primarily via the following objectives:

- to give effect to and to regulate the fundamental rights contained in Section 27 of the constitution;
- to give effect to the duties of the Republic as a member state of the International Labour Organisation;

- to provide a framework in which employees and their unions, employers and employer associations can bargain collectively to determine wages, terms and conditions of employment and other matters of mutual interest, and formulate industry policy; and

- to promote:
 - ◆ orderly collective bargaining,
 - ◆ collective bargaining at sectoral level,
 - ◆ workers' participation and decision-making at the workplace, and
 - ◆ the effective resolution of disputes (Government Gazette, 1995, pp. 1-3).

2.3.5.2 Legislation on valid recruitment and selection procedures

Organisations have to ensure that they comply with the LRA by using valid and fair recruitment and selection procedures. Failure to comply with these legal requirements will result in what is known as unfair labour practice. Unfair labour practice is defined as “any unfair practice or omission which arises between an employer and employee” (Bendix, 1996, p.269).

For the purpose of unfair discrimination, an applicant for a position may be regarded as an employee. An employer is not prevented from adopting a policy or practice aimed at the protection and advancement of employees previously disadvantaged by unfair discrimination or from appointing persons in terms of the inherent requirements of a job (Bendix, 1996).

Whenever unfair discrimination is alleged, the onus is on the employer to establish that a specific practice is fair. The Employment Equity Act 56 (1998) also states that psychological testing is prohibited, unless it:

- Is scientifically valid and reliable;
- Is applied fairly to all employees; and

- Is not biased against any employee or group.

2.3.5.3 *Guidelines for assessment practices*

Employers and organisations should be sensitive about issues regarding bias and fairness, given South Africa's diverse multi-cultural context. The Society for Industrial Psychology (SIP) has proposed a Code of Practice for psychological assessment in an attempt to promote fairness in the workplace (Code of Practice for Psychological Assessment in the Workplace, 1998). The code proposes the following guidelines:

- assessment practitioners should ensure that assessment methods are not used with people for whom the method is not appropriate;
- assessment practitioners should be aware of the impact on assessment of cultural, linguistic and disability factors, and of aspects of disadvantage;
- wherever possible, the potential impact of bias should be reduced by using a range of methods that vary in terms of constructs, format and time pressure;
- it is professionally responsible to conduct research or make data available for research on the bias and validity of assessment method, and make the results available beyond the assessment practitioners' organisation;
- assessment practitioners, and psychologists in particular, should have a thorough understanding of the various fairness models and should advise stakeholders of their advantages and disadvantages;
- these models apply to all assessment methodologies, as all methods (including those not recognised as psychological tests) are subject to bias.

2.4 THE PROCEDURE OF VALIDATING A SELECTION BATTERY

In selecting or promoting employees, the practitioner needs an answer to a basic question: Do candidates who perform better on this test, perform better on the job? The best way of answering this question is by conducting a validation study. The procedure of developing a new selection battery and validating an existing one is basically similar. There are however some differences. In order to highlight these differences, the two procedures are discussed separately. The procedure of developing a scientific personnel selection battery has already been discussed in paragraph 2.2.

According to Klinvex (1999) the major steps in conducting a criterion-related validation study are the following:

2.4.1 Job analysis

Job analysis refers to the systematic study of job content and job context for the purpose of obtaining a detailed statement of work behaviours and other information relevant to the job. In test validation, the purpose of job analysis is to identify those aspects of the job that will serve as the criteria of job performance to be "predicted" by the tests and to identify the appropriate selection instruments that will make up the trial test battery.

2.4.2 Development of criterion measures of job performance

Criterion development is arguably the most important step in the validation process since criterion measures should represent those aspects of worker behaviour that are relevant to the organisation's core business and that validated tests seek to predict. Typical criterion measures are production data, personnel data and supervisory evaluations.

2.4.3 Selection of predictors

The term “predictor” refers to the selection instrument that is validated for the purpose of determining whether the skill, ability or worker characteristic being measured by the selection instrument is correlated with performance on the criterion. Examples of predictors are, skill or ability tests, personality or interest inventories, knowledge tests, interviews and reference checks.

2.4.4 Composition of study sample

The sample in a criterion-related validation study refers to those individuals to whom the experimental battery of tests will be administered and whose on-the-job performance will be used as criterion measures. Kline (1999) further maintains that two conditions may render a criterion-related study technically infeasible. One constraint has to do with severe restriction of range on either the predictor or the criterion variable. The other condition deals with sample size. To be effective, validation studies require testing a fairly large number of individuals (between 60 – 100 or more). Thus, adequate variability in predictor and criterion scores as well as adequate sample size are threshold requirements for a criterion-related validation study.

2.4.5 Statistical analysis

Statistics play three general roles in a criterion-related validation study:

- a. To summarise the data for ease of understanding. The relationship between test scores and criterion scores, is expressed by the correlation coefficient.
- b. To “infer” by evaluating whether obtained results are statistically significant or whether they can be attributed to chance.

- c. To assemble the optimal battery of tests for operational use. The interest is in determining which tests are to be used in combination and how each test is to be weighted.

2.4.6 Implementation of validity study results

Since validation studies could be an expensive exercise, it is important to use the results to the best advantage of the organisation. It has already been indicated that a thorough job analysis is suitable not only for test validation purposes, but for training and job evaluation as well. A validation study should not be regarded as an isolated exercise, but as an integral component of the entire human resource function.

A common discovery in cross validation research is that a test predicts the relevant criterion less accurately with the new sample of examinees than with the original sample. The term *validity shrinkage* is applied to this phenomenon (Gregory, 1996). Validity shrinkage is an inevitable part of test development and underscores the need for cross-validation. In most cases, shrinkage is slight and the instrument withstands the challenge of cross-validation. However, shrinkage of test validity can be a major problem when derivation and cross-validation samples are small, the number of potential test items is large, and items are chosen on a purely empirical basis without theoretical rationale (Anastasi & Urbina, 1997; Gregory, 1996).

2.5 THE PSYCHOLOGICAL PROFILE OF A PILOT IN THE SAAF

In the most recent publication that was produced to serve as guideline for the selection of pilots in the SAAF for the 1990s, Aspeling (1990) proposed that a pilot should have the following critical abilities, characteristics and personality requirements (see Table 2.1). The tests currently included in the selection battery to measure specific abilities and characteristics and which characteristics they measure are also indicated in Table 2.1.

TABLE 2.1 PSYCHOLOGICAL PROFILE OF PILOTS IN THE SAAF AS MEASURED BY SELECTION BATTERY

PROFILE CHARACTERISTICS	BLOX TEST	ADVANCED RAVENS TEST	VIENNA DETERMINATION TEST
Primary cognitive abilities			
Mental alertness	*	*	*
Memory		*	*
Technical comprehension	*		
Technical knowledge	*		
Arithmetic ability		*	
Initiative			
Primary affective characteristics			
Stress management			*
Adaptability			*
Emotional stability			
Assertiveness			
Institutionalised aggression			
Caution			*
Primary psycho-physiological characteristics			
Vitality			
Man-machine adaptability			*
Alertness	*	*	*
Concentration	*	*	*

PROFILE CHARACTERISTICS	BLOX TEST	ADVANCED RAVENS TEST	VIENNA DETERMINATION TEST
Primary conative characteristics			
Motivation			*
Responsibility			
Primary psycho-social characteristics			
Leadership			
Team Work			
Discipline			
Interpersonal relations			
Communication			
Primary perceptual-motor characteristics			
Complex co-ordination	*		*
Reaction time			*
Dynamic anticipation	*	*	*
Primary management ability			
Mature decision-making			
Strategic planning			

An overview of the profile presented above, indicates that a large number of characteristics are not being measured by the current trainee pilot selection battery. There appears to be an emphasis on cognitive abilities and perceptual-motor functioning.

Characteristics such as initiative, emotional stability, assertiveness, sense of responsibility and other psycho-social traits (inter-personal relations, communication and self-discipline) are inadequately assessed, or not at all. It is recommended that a personality test as well as simulation exercises be included to assess the above-mentioned characteristics.

2.6 CONCLUSION

In our current competitive environment and stringent labour legislation, scientific selection is not only necessary, but imperative. The purpose of scientific personnel selection is to identify those candidates with the required skills, knowledge and aptitudes for the successful execution of a specific job. In order to have access to such a selection battery, research should be undertaken to design a new battery or to evaluate and validate an existing one. Hence, the significance of this research.

In the following chapter the method to be utilised for the empirical investigation is discussed.

CHAPTER 3

EMPIRICAL STUDY

This chapter presents a detailed discussion of the method employed for the empirical study. Aspects to be discussed are the steps that were followed in order to guide the investigation, predictor and criterion variables, the statistical techniques used to analyse the data and research hypotheses.

3.1 POPULATION AND SAMPLE

The population comprised all candidates who have been selected for pilot training in the SAAF between 1997-1999, and the sample comprised those individuals who have completed the Ground School Phase (18 months) of the trainee pilot training programme. A sample of convenience was used, since only information available on database could be used for the empirical study.

The sample consisted of 92 candidates who have successfully completed the Ground School Phase of the trainee pilot training programme between 1997 and 1999. Unfortunately, due to the nature of the training and the resources available for pilot training, only a small number of candidates can be trained during any given training cycle. This together with the fact that the 2000 – 2001 candidates were still in the process of being trained at the time the study was conducted, limited the sample size. Table 3.1 is an indication of the distribution of candidates over three years of training.

TABLE 3.1 DISTRIBUTION OF SAMPLE BY YEARS OF TRAINING

Year	Frequency	Percentage	Cumulative Percentage
1997	28	30.4	30.4
1998	27	29.3	59.8
1999	37	40.2	100.0
TOTAL	92	100.0	

In terms of gender the sample was represented as follows.

TABLE 3.2 COMPOSITION OF THE SAMPLE ACCORDING TO GENDER

Gender	Frequency	Percentage	Cumulative Percentage
Male	84	91.3	91.3
Female	8	8.7	100.0
TOTAL	92	100.0	

The sample is disproportionate in terms of gender, which is to be expected as pilot training has always been regarded as a male dominated profession. However, this situation (perception) is changing.

The sample consisted of the following race groups.

TABLE 3.3 COMPOSITION OF THE SAMPLE ACCORDING TO RACE GROUPS

Race groups	Frequency	Percentage	Cumulative Percentage
White	74	80.4	80.4
Black	7	7.6	88.0
Asian	4	4.3	92.4
Coloured	7	7.6	100.0
TOTAL	92	100.0	

In terms of home language the sample was represented as follows.

TABLE 3.4 COMPOSITION OF THE SAMPLE ACCORDING TO HOME LANGUAGE

Home language	Frequency	Percentage
Afrikaans	49	53.3%
English	37	40.2%
Sesotho	5	5.4%
IsiZulu	1	1.1%
TOTAL	92	100%

The small number of candidates speaking Sesotho and Isizulu could be a reflection of the socio-political situation in the past, in the sense that very few black candidates regarded pilot training as a viable career to be followed.

The demographics of the South African population indicate that the sample is not representative in terms of home language, race or gender. This limitation could not be corrected since the researcher used all the data available on database for the empirical study. It is envisaged that the representivity of the groups in terms of race and gender will improve as the pool of, for example, black candidates and women is enlarged.

3.2 MEASURING INSTRUMENTS

The current pilot selection battery consists of the Blox test, Raven's Advanced Progressive Matrices test and the Vienna Determination test. These instruments serve as the independent variables of this study. What follows is the aim, description, administration and the confirmed reliability and validity of these instruments (Holburn, 1992; Raven & Court, 1985; Schufried, 1997).

3.2.1 The Blox Test

The aims, description, administration, reliability and validity of the Blox test are presented below.

3.2.1.1 *The aim of the test*

The aim of the Blox test is to measure spatial ability, which is made up of a number of factors, namely, spatial relations, spatial orientation and visualisation.

3.2.1.2 *Description of the test*

The Blox test is a paper and pencil test bound in a re-usable book. This non-verbal test consists of 6 examples and 45 test items. The stimulus consists of isometric drawings of different combinations of two, three, four, five or six cubes. Each set of cubes must be compared to similar arrangements of cubes viewed from other angles. The first factor (spatial relations and orientation) has been described as the ability to comprehend the nature of the arrangements within a visual stimulus pattern primarily with respect to the examiner's body or frame of reference (Zimmerman, 1957). The parts that comprise the figure, maintain their relationships to one another as the whole figure is moved or rotated in space. The candidate's task would be to recognise the same visual stimulus pattern from different angles.

The second factor is visualisation, which involves the ability to mentally manipulate (rotate, invert, twist) one or more parts of a visual stimulus pattern and to recognise the changed appearance of the object. Visualisation involves "movement among the internal parts" of an object (McGee, 1979). Both the above abilities correlate with a higher level of technical ability.

3.2.1.3 *Administration of the test*

The test administrator goes through the six test examples with the candidates in order to ensure that they have a clear understanding of what is expected. The candidates must analyse each stimulus set and choose the corresponding set seen from a different angle, from the five possible responses. The time limit for the test is 30 minutes (Holburn, 1992).

3.2.1.4 *Validity and reliability of the test*

The construct and predictive validity of the test was confirmed in a study which was conducted on trainee engineering technicians (Holburn, 1992). Internal consistency data indicated correlations of consistency of 0,91 and reliability of at least 0,81 (Holburn, 1992).

3.2.2 **The Advanced Progressive Matrices Test**

The aims, description, administration, reliability and validity of the test are presented below.

3.2.2.1 *The aim of the test*

The Advanced Progressive Matrices test (APM) provides a means of assessing more accurately a person's speed of intellectual work. By imposing a time limit it can be used to assess a person's "intellectual efficiency" in the sense of his/her present speed of accurate intellectual work. This is generally related to a person's total capacity for orderly thinking (Raven & Court, 1985).

3.2.2.2 *Description of the test*

Set 1 consists of only 12 problems, which is followed immediately by set 2. Set 2 consists of 36 problems, arranged in ascending order of difficulty. Thus it is possible

that not every candidate will attempt every problem before stopping. The items consist of a number of designs arranged in rows and columns, from each of which a part has been removed. Respondents are presented with test items in the same sequence and instructed to proceed as quickly as possible (Raven & Court, 1985).

3.2.2.3 Administration of the test

Set 1 is used to provide the necessary training in the method of working. This is followed immediately by set 2. Set 2 is the actual test. The task is to identify, from given alternatives, the missing part of the design. Respondents must indicate the missing part of the design by making a cross over the number corresponding with the correct alternative on the answer sheet provided. The time restriction for the completion of the test is 45 minutes (Raven & Court, 1985).

Detailed instructions for marking, evaluating and reporting of results are provided in the Manual for Advanced Progressive Matrices and Vocabulary scales (Raven & Court, 1985).

3.2.2.4 Validity and reliability of the test

In terms of validity studies, Raven and Court (1985) report that the majority of studies giving internal consistency data report correlations of consistency of at least 0,90 with a modal value of 0,91. The Raven's APM has also been found to correlate 0,74 with the full Wechsler Adult Intelligence Scale and 0,75 with the Otis I.Q test (Raven & Court, 1985). Raven and Court (1985) also report that substantial studies (not described in the source) have indicated that a reliability of at least 0,86 can be obtained for Raven's APM. Burke (in Raven & Court, 1985) summarised consistency measures for Raven's APM using 567 subjects and found correlations of up to 0,95 for subjects aged 56-65 and up to 0,83 with younger adults. Irvine (in Raven & Court, 1985; Verguts & De Boeck, 1997) is reported to have obtained a correlation of 0,90 for a sample of 1600 Rhodesian school children.

3.2.3 The Vienna Determination Test

The aims, description, administration, reliability and validity of the Vienna Determination test are presented below.

3.2.3.1 *The aim of the test*

The Vienna Determination test is a general measure for “reactive capacity” and reaction speed (Schuhfried, 1997). The aim of the procedure is to measure the following *cognitive performances*:

- discrimination between colours and tones;
- memorising of relevant aspects of the stimulus configuration and the response buttons and the relation between them; and
- selection of the relevant reaction as learned in the instruction or during the test run.

In addition, *motor performance* is measured by pressing the keys and foot pedals. As there is a simple relation between stimulus and reaction the main stress lies in the permanent and continuous quick reaction to changing stimuli. The challenge of the Vienna Determination test depends on two variables: the speed of stimulus presentation and the number of stimuli and reactions (Schuhfried, 1997).

3.2.3.2 *Description of the test*

The Vienna Determination test is a complex multi stimuli reaction test. The stimuli are presented on the screen and the response is done by pressing the corresponding button on the universal panel. Optical signals and tone have to be responded to by pressing buttons or foot pedals. Kissler (1986, p.226) is of the opinion that “with such units it is possible to measure behaviour under high psychological and physiological stress, because the high frequency of the signals

lead to an overcharge situation for almost everybody. Insufficient response behaviour is the result".

The stimuli are presented on the screen. Ten optical stimuli with the colours white, yellow, red, green and blue are sorted in two lines. The universal panel has five buttons in the corresponding colours and enables the candidate to work with both hands. Two stimuli (white, rectangular fields on the monitor) require a reaction with the left or right foot pedal. Furthermore, there are two acoustical signals (high and low tone) which are responded to by pressing the grey (upper) or black (lower) bar button.

There are three kinds of stimulus presentation depending on the time limit:

- 1) *Modus reaction* (fixed presentation time); Each stimuli is presented for a fixed time, then the next follows whether a response has been given or not;
- 2) *Modus action* (free presentation time) a new stimulus is presented after the response to the last one is done, so the presentation time is determined by the candidate;
- 3) *Modus adaptive* (varying presentation time) the frequency of stimulus presentation is controlled by the working speed of the candidate (Schuhfried, 1997).

3.2.3.3 Administration of the test

The candidate goes through a practice phase first in modus action and then in modus reaction, in order to become familiar with the various stimuli, for example, different colours, lights, foot pedals and acoustical signals. If more than three errors are made or no entry is made within 45 seconds for three times the practising phase is aborted automatically. The candidate is asked to inform the test administrator to help the candidate through the instruction. The test phase starts immediately after

the practising phase, which consists of three distinct phases. The entire system is computer operated. On completion of the third phase, a report is generated by the computer (Schufried, 1997).

To keep the test administrator informed of every step in the testing procedure the programme will display the following information on the test administrator's screen:

- ◆ Name of test and version number
- ◆ Number of parameter block and identification
- ◆ Modus
- ◆ Length of test
- ◆ Remaining time
- ◆ Current stimulus length
- ◆ Number of presented stimuli
- ◆ Number of correct responses, and
- ◆ median of reaction time (Schufried, 1997)

3.2.3.4 *Validity and reliability of the test*

Test results indicate (detailed tables are presented in the Vienna Test System Manual, Schufried, 1997) that the consistency with which the single variables are measured is generally extremely high. This means that the Vienna Determination Test is a diagnostic instrument of particularly high precision.

In the frame of a standardisation study (Weinkirn, 1996, cited in Schufried, 1997) standard parameter block S1 has been presented together with the Reaction Test (RT) and Tachistoscopic Traffic Test Mannheim for Screen (TAVTMB) to a population of 180 healthy subjects. There were high statistically significant correlations between the two MEDIANS OF REACTION TIMES (MD.RT). The TOTAL OF CORRECT show high significant correlation with the number of correct answers of the TAVTMB. The TAVTMB is a procedure which examines the optical

perceptual speed for special traffic situations presented as tachistoscopic pictures (Schufried, 1997).

For the purpose of this study, the principle components of the current pilot selection battery (the Blox test, the Advanced Progressive Matrices test, and the Vienna Determination test) serve as predictor variables. The Military Psychological Institute (MPI) is responsible for the pre-selection of all potential candidates. Their psychological results are available on database at MPI.

3.3 DATA COLLECTION PROCEDURE

The data collection procedures in respect of the independent and dependent variables are presented below.

3.3.1 Data collection in respect of independent variables

The psychometric tests were administered and scored by Psychometrists and Psychologists at MPI under standard testing conditions. The Blox test and the Ravens APM test were administered in paper and pencil format in a single testing session. The Vienna Determination Test was administered on the following day. The item responses of each candidate on each test were captured, as were the raw scores obtained on each of the tests. Scoring masks obtained from the test publishers were used to score all the tests, except for the Vienna Determination test. In the case of the Vienna Determination test, a computer is used to generate a report that contains the test results of the candidates.

3.3.2 Data collection in respect of dependent variable

At the end of the Ground School Phase of the trainees' training, standard written examinations are administered to each candidate. The final Ground School Phase results of those candidates who have completed their training, was used as dependent variable in this study. This procedure had certain implications in terms of

restriction of range, which is highlighted as one of the limitations of the study. The total percentage score of each candidate was used as criterion variable.

3.4 DATA PROCESSING

Data was processed by using the Statistical Package for Social Science (SPSS) software programme. The data was analysed with SPSS using the following techniques (SPSS Base 9.0, 1999):

- a. Pearson Product–Moment Correlations
- b. Multiple Linear Regression Analysis

The purpose of statistical analysis is to test the hypotheses (presented in 3.5), that the psychometric results of the respondents on the different psychometric tests that compose the selection battery, correlate significantly with the same respondents' Ground School Phase results.

The following statistical techniques and procedures were used:

3.4.1 Correlation coefficient

One general way of expressing the relationship between scores on a psychological test and a measure of job behaviour is by means of the correlation coefficient. This is a single index, varying from -1, 00 to + 1, 00 which expresses the direction and degree of the amount of co-variation between two sets of scores (Kerlinger, 1986).

According to De la Rey (1978) the Pearson Product Moment correlation coefficient should reach a specific magnitude before the correlation can be regarded as a statistically valid relationship. The strength of the correlation coefficient is also determined by the number of participants in any given study.

The correlation coefficient is often used for the following reasons (Huysamen, 1980):

- a) Correlation techniques are familiar to most psychology practitioners, being part of their training;
- b) the correlation coefficient is a convenient way of summarising a relationship into one general descriptive term; and
- c) the actual correlation coefficient obtained, the value of the statistic, is directly convertible into a measure of predictive accuracy, by squaring the obtained correlation.

Correlations between the dependent and independent variables are interpreted as follows: a correlation between 0,1 – 0,2 is regarded as weak, a correlation between 0,3 – 0,4 mediocre and correlation above 0,5 as strong (Howell, 1995).

Through statistical inference the “*statistical significance*” of certain results can be determined. It indicates the chances or likelihood of a certain result being due to chance errors. Generally, in practice, if the probability is less than 5 out of 100 that a certain result can be ascribed to the operation of chance factors, the result is said to be “statistically significant” (Gregory, 1996).

According to Tabachnick and Fidell (1996), one needs to compare the correlations of the independent variables with each other, the unique relationship of each of the independent variables with the dependent variable and the total relationship of all the independent variables with the dependent variable in order to get a complete picture of the function of an independent variable.

Three aspects are therefore important for this analysis:

- a. Independent Variable correlations: Correlations of the independent variables with each other;

- b. The unique relationship of independent variables with the dependent variable: Individual R square values for each sub test (bivariate regression analysis); and
- c. The total relationship of independent variables with the dependent variable: R square for all the independent variables together as a group (by means of multivariate regression analysis with “Enter” method).

The results are reported in the above-mentioned format. The tables (in chapter 4) can be interpreted as follows:

- a. **R** stands for the Pearson correlation between the independent variables (Blox, Advanced Ravens and Vienna Determination test) and dependent variable (Ground School Phase results).
- b. **R square** indicates the percentage of the variance in the dependent variable explained by the independent variable(s). (Multiply R square with 100 to obtain the percentage variance explained by each sub-test).
- c. **Adjusted R square** is an adjustment for the fact that when one has a large number of independent variables, it is possible that R square will become artificially high simply because some independent variables’ (sub tests) chance variations “explain” small parts of the variance of the dependent variable. At the extreme, when there are as many independent variables as cases in the sample, R square will always be 1,0 (Anastasi & Urbina, 1997; Tabachnick & Fidell, 1996).

When used for the case of few independent variables, R square and adjusted R square will be close. When there are many independent variables, adjusted R square may be noticeably lower. The greater the number of independent variables, the more the researcher is expected to report the adjusted

coefficient, hence the report on adjusted R square for this analysis (despite the fact that three independent variables are not that many).

- d. The **Beta** weights can be compared to judge relative predictive power of independent variables (sub tests), i.e., determine which tests are the better predictors of performance.

3.4.2 Multiple regression

This is a statistical technique indicating the maximum predictive validity obtainable from the combination of scores on various predictors. It refers to the procedure for combining predictors in such a way as to yield the highest multiple correlation (**R**) between the set of predictors and the criterion of job performance (Cascio, 1991).

The multiple regression approach to combining predictor information has four main advantages (Cascio, 1991):

- a) It minimizes errors in prediction by ensuring that predictors are combined to obtain the most efficient (highest) estimate of subsequent performance;
- b) It is a highly flexible approach, because it makes possible the setting-up of equations for each of a number of jobs using the same predictors;
- c) The flexibility of the approach extends to allowing the organisation to be either applicant-centred (each person is placed where their predicted score is highest even if this is below that required for success) or organisation-centred (each person is placed in that job where their predicted score is furthest above the minimum score required for success);
- d) The multiple regression approach makes it possible for high score in some predictors to compensate for low scores in other predictors. This is an important point for it is often the case that substitution of one skill or ability can compensate for deficiencies in another (Jacobs, Haasbroek & Theron, 1992).

However, this could also be a problem. One should check for critical minimum performance levels in specific predictors.

Linear correlations between individual independent variables and the dependent variable were also determined. The rationale behind and utilising this technique can be delineated as follows (Tabachnick & Fidell, 1996):

- a. To determine whether there is a linear relationship between the independent variables (individually) with the dependent variable.
- b. No shared variance has been declared.

In the standard, or simultaneous, model all the independent variables enter into the regression equation at once; each one is assessed as if it had entered the regression after all other independent variables had entered. Each independent variable is evaluated in terms of what it adds to prediction of the dependent variable that is different from the predictability afforded by all the other independent variables.

In standard multiple regression, it is possible for a variable to appear unimportant in the solution when it actually is highly correlated with the dependent variable (Tabachnick & Fidell, 1996). If the area of that correlation is whittled away by other independent variables (like overlapping circles), the unique contribution of the independent variable is often very small despite a substantial correlation with the dependent variable. For this reason, both the full correlation and the unique contribution of the independent variables need to be considered in interpretation (Tabachnick & Fidell, 1996). Hence, the importance of this technique in this study.

3.5 RESEARCH HYPOTHESES

Hypotheses in respect of the validity of each psychological test for the prediction of training success are provided below:

- ◆ H0 Raven's Advanced Progressive Matrices is not statistically significant for the prediction of training success of trainee pilots during the Ground School Phase
- ◆ H1 Raven's Advanced Progressive Matrices is statistically significant for the prediction of training success of trainee pilots during the Ground School Phase
- ◆ H0 The Blox test is not statistically significant for the prediction of training success of trainee pilots during the Ground School Phase
- ◆ H1 The Blox test is statistically significant for the prediction of training success of trainee pilots during the Ground School Phase
- ◆ H0 The Vienna Determination test is not statistically significant for the prediction of training success of trainee pilots during the Ground School Phase
- ◆ H1 The Vienna Determination test is statistically significant for the prediction of training success of trainee pilots during the Ground School Phase
- ◆ H0 The total selection battery is not statistically significant for the prediction of training success of trainee pilots during the Ground School Phase
- ◆ H1 The total selection battery is statistically significant for the prediction of training success of trainee pilots during the Ground School Phase

3.6 CONCLUSION

This chapter outlined the method of investigation that was used for the empirical study. The objects of the study, namely, the predictor and criterion variables were consequently discussed at length. The statistical processing of data collected is discussed in terms of the various techniques used to determine the predictive validity

of each of the psychological tests. The chapter concludes by outlining the statistical techniques that were used for the analysis of the data and the research hypotheses posed in respect of the validity of each of the psychological tests used for the prediction of training success during the Ground School Phase.

The following chapter deals with the reporting of the data and the interpretation of the results.

CHAPTER 4

RESULTS

In this chapter, the results of the empirical study are reported. Results are presented in respect of the validity of the current selection battery for the prediction of training success. The results are interpreted within the context of the empirical study and the literature review. The results will also provide the basis for the rejection or confirmation of the research hypotheses in chapter three.

4.1 RESULTS OF INDEPENDENT VARIABLE CORRELATIONS

The primary aim of this section is the following:

- a. To report on the strength of the relationship between the dependent and independent variables.
- b. To screen for extremes, multicollinearity and singularity. Multicollinearity and singularity are problems with a correlation matrix that occur when variables are too highly correlated. With multicollinearity, the variables are very highly correlated (say 0,90 and above); with singularity, the variables are redundant; one of the variables is a combination of two or more of the other variables. It is evident that it could cause both logical and statistical problems. No indication of multicollinearity or singularity were found.
- c. To determine which independent variables correlate strongly with the dependent variable.

The interdependent correlations between the independent variables are reported in table 4.1, and the correlations between the independent variables and dependent variable are reported in table 4.2.

TABLE 4.1 INTERDEPENDENT CORRELATIONS BETWEEN INDEPENDENT VARIABLES

		Blox Scores	Adv Ravens scores	Vienna 1	Vienna 2	Vienna 3
Blox scores	Correlation	1	0,526	0,004	0,136	0,075
	Sig.2 Tail	.	0,000	0,969	0,196	0,479
	N	92	92	92	92	92
Adv Ravens Scores	Correlation	0,526**	1	0,141	0,191	0,198
	Sig.2 Tail	0,000	.	0,179	0,068	0,059
	N	92	92	92	92	92
Vienna 1	Correlation	0,004	0,141	1	0,378**	0,378**
	Sig.2 Tail	0,969	0,179	.	0,000	0,000
	N	92	92	92	92	92
Vienna 2	Correlation	0,136	0,091	0,378**	1	0,563**
	Sig.2 Tail	0,196	0,068	0,000	.	0,000
	N	92	92	92	92	92
Vienna 3	Correlation	0,075	0,198	0,378**	0,563**	1
	Sig.2 Tail	0,479	0,059	0,000	0,000	.
	N	92	92	92	92	92

** Correlation is significant at the 0,01 level (2-tailed)

TABLE 4.2 PEARSON CORRELATIONS BETWEEN INDEPENDENT VARIABLES AND DEPENDENT VARIABLE

	Blox	Advanced Ravens	Vienna 1	Vienna 2	Vienna 3
Results for Ground School	-0,091	0,275**	0,171	0,186	0,344**
P-values	0,032N/S	0,008	0,103N/S	0,076N/S	0,028

** Significant on 1%

* Significant on 5%

N/S = Not significant

Table 4.2 indicates that the Advanced Ravens test and the Vienna 3 (Third phase) reflect a statistically significant positive correlation. The Vienna Phases 1 and 2 reflected no statistically significant relationship. These two phases of the Vienna Determination test are therefore, not good predictors of training success during the Ground School Phase.

The results of the correlation analysis can be summarised as follows:

- a. The correlation between the Blox test and the Ground School Phase results is not statistically significant ($r = 0,091$; $p > 0,05$).
- b. The Advanced Ravens test has a statistically significant correlation with the Ground School Phase results ($r = 0,275$; $p \leq 0,01$), which meets the critical cut-off point of this study, namely, ($p \leq 0,05$).
- c. The Vienna test phases 1 and 2 yielded statistically insignificant results ($p > 0,05$), i.e. 0,171 and 0,186 respectively.

- d. The Vienna test phase 3 has a weak positive correlation with the Ground School Phase results, i.e. ($r = 0,048$; $p < 0,05$).

Possible reasons for these results could be found by taking a closer look at the nature of the Vienna Determination test.

The three phases of the Vienna test present three kinds of stimuli, depending on the time limit.

Modus reaction (phase one) has a fixed presentation time. Each stimuli is presented for a fixed time, then the next follows whether a response has been given or not.

Modus action (phase two) has a free presentation time. A new stimulus is presented after the response to the last one is done, so the presentation time is determined by the candidate.

Modus adaptive (phase three) has varying presentation time. The frequency of stimulus presentation is controlled by the working speed of the candidate.

Possible reasons why the Vienna Phases 1 and 2 had no statistically significant relationship with the dependent variable could be the following:

- a. Since the Vienna phase 1 has a fixed presentation time, most candidates get into the "rhythm" of the test quite quickly. This results in them scoring above 70% for this phase.
- b. Since Vienna phase 2 has a free presentation time (the presentation time is regulated by the candidate), there is little pressure on the candidate, resulting in them scoring above 80% for this phase. What has been reported in the above two paragraphs, leads to restriction of range, which then probably lowers the correlation coefficient.

- c. The Vienna phase 3 is a better predictor than the first two phases of the test. This could be attributed to the fact that the Vienna phase 3 presents the candidate with a variety of stimuli, resulting in significantly higher stress levels on the part of the candidate. This is a much better reflection of typical flying conditions the candidate will be exposed to, as opposed to the first two phases of the test.

The Advanced Ravens test provides an indication of a person's intellectual efficiency (e.g. to make quick, accurate judgements under stressful conditions) which is a critical skill during the flying process. Hence, the positive correlation between this test and the Ground School Phase results (dependent variable) of the candidates.

Despite the Blox test's poor correlation with the Ground School Phase results, it remains an important component of the selection battery in the sense that it measures one of the critical abilities a potential pilot should have, namely, spatial relations and orientation. It should be retained until a more comprehensive study is undertaken.

4.2 RESULTS OF THE TOTAL RELATIONSHIP OF INDEPENDENT VARIABLES AND DEPENDENT VARIABLES

Table 4.3 is a summary of the total relationship of the independent variables (Blox, Advanced Ravens and the Vienna test) with the dependent variable (Ground School Phase results) and table 4.4 is an indication of the individual beta weights.

TABLE 4.3 PREDICTIVE VALUES OF INDEPENDENT VARIABLES WITH DEPENDENT VARIABLES

	R	R-squared	Adjusted R-squared	Significance Level
Total Battery	0,485	0,235	0,191	0,000

The total selection battery has a significant linear relationship with the dependent variable, i.e. (0,485). The relationship between the independent variables as a whole and the dependent variable is therefore, relatively strong. The total selection battery explains 19.1% of the variance in the dependent variable. Only a small part of the variance of the dependent variable is explained by the **Adjusted R square**.

TABLE 4.4 INDIVIDUAL BETA WEIGHTS

	Beta	Significance Level
Blox	-0,091	N/S
Advanced Ravens	0,382**	1 %
Vienna 1	0,012	N/S
Vienna 2	- 0,016	N/S
Vienna 3	0,296*	5 %

** Significant on 1% level

* Significant on 5% level

N/S = Not significant

The beta weights provide an indication of the relative predictive power of the independent variables. Table 4.4 indicates that the Advanced Ravens test, the Blox test and the Vienna Determination test (Phase 3) have the strongest predictive power.

The total selection battery has a significant linear relationship with the dependent variable (see table 4.3). The **R** of 0,485 indicates that the relationship between independent variables and the dependent variable is relatively strong and statistically highly significant. The independent variables (Blox test, Advanced Ravens test and the Vienna Determination test) collectively explain 23,5% of the variance in the dependent variable i.e. the Ground School Phase results (R Square = 0,235). The selection battery declares 19,1% of the variance in the dependent variable. The

Adjusted R square thus “explains” only a relatively small part of the variance of the dependent variable.

The beta weights can be compared to judge relative predictive power of independent variables, i.e. determine which tests are the better predictors of performance. By examining these weights, one will notice that they reflect the relative strengths of each independent variable.

The beta weights (see table 4.4) indicate that the Advanced Ravens test is the strongest predictor of Ground School Phase results (beta = 0,382; $p < 0,01$), followed by the Blox test (beta = -0,091; $p < 0,01$) and finally, the Vienna Determination test Phase 3 (beta = 0,296; $p < 0,05$).

The beta weights also indicate that the Vienna Test Phases 1 and 2 have no statistically significant linear correlation with the dependent variable i.e. Ground School Phase results.

4.3 RESULTS OF INDIVIDUAL RELATIONSHIPS OF INDEPENDENT VARIABLES AND DEPENDENT VARIABLES

The purpose of this analysis is to determine whether there is a linear correlation between individual independent variables and the dependent variable when considering the Adjusted R-squared value. These results are reported in table 4.5.

TABLE 4.5 CORRELATION BETWEEN INDEPENDENT VARIABLES AND DEPENDENT VARIABLE

	R	R squared	Adjusted R squared	Beta	Significance Level
Blox	-0,091	0,008	-0,003	-0,091	N/S
Adv Ravens	0,275	0,076	0,066	0,275	0,008
Vienna 1	0,171	0,029	0,018	0,171	(N/S)
Vienna 2	0,186	0,035	0,024	0,186	(N/S)
Vienna 3	0,344	0,119	0,109	0,344	0,001

The results as reported in table 4.5, primarily indicate a statistically significant linear relationship between the Advanced Ravens test and the criterion (the Ground School Phase results of the candidates). The Vienna Determination test phase 3 explains 10% of the variance in the dependent variable. The rest of the test battery as reported earlier, failed to indicate a statistically significant relationship with the dependent variable.

As reported in paragraph 3.4.2, linear correlations between individual independent variables and the dependent variable were also determined.

The rationale behind and utilising this technique are:

- a. To determine whether there is a linear relationship between the independent variables (individually) with the dependent variable.
- b. No shared variance has been declared.

The unique relationship of each of the independent variables with the dependent variable yielded the following results:

- a. There is a statistically significant linear relationship between the Advanced Ravens test and the Ground School Phase results of the candidates. It explains 6,6% of the unique variance of the dependent variable (Adjusted R-squared = 0,066; $p < 0,01$).
- b. The unique variance of the Vienna Determination test phase 3 explains 10% of the variance in the dependent variable (Adjusted R-squared = 0,109; $p < 0,01$).
- c. On their own, the rest of the tests in the battery i.e. the Blox test and the Vienna Determination test phases 1 and 2, failed to indicate a statistically significant relationship with the dependent variable.

Possible reasons for the results reported above have already been discussed in paragraph 4.1.

4.4 INTEGRATION OF RESULTS

The overall results of this study seem to indicate that the current trainee pilot selection battery has some utility value. Some of the major conclusions that can be drawn from the results of the study are the following:

- a. The total selection battery has a significant linear relationship with the dependent variable ($R=0,485$, $p=0,000$). (See table 4.3)
- b. The Advanced Ravens ($r = 0,275$) and the Vienna Determination test phase 3 ($r = 0,344$) correlate better than the other instruments with the Ground School Phase results. (See table 4.5).
- c. The Vienna Determination test phase 1 and 2 yielded statistically insignificant results, i.e. 0,171 and 0,186 respectively.

- d. The independent variables collectively explain approximately 19,1% of the variance in the dependent variable i.e. Adjusted R-squared = 0,191.
- e. The beta weights confirm that the Advanced Ravens test is the strongest predictor of Ground School Phase results, followed by the Blox test and the Vienna Determination test phase 3.

The significance levels (p-values) provide an indication of whether the null-hypotheses can be rejected or not. A significance level of less than 0,01 or 0,05 facilitates the rejection of the null-hypothesis (Anastasi & Urbina, 1997; Huysamen, 1980; Rosnow & Rosenthal, 1996).

Based on the results of this study, the null hypothesis regarding the predictive power of the Advanced Ravens test and the Vienna Determination test can be rejected. However, on the basis of the results of this study, the null hypothesis regarding the predictive power of the Blox test cannot be rejected with confidence. Based on the results of this investigation, the null hypothesis regarding the predictive power of the total selection battery can be rejected. The battery can be adapted to increase its predictive validity, in the light of the statistical significance of the Blox test and the Vienna Determination test phases 1 and 2.

When the current profile (see figure 2.1) is compared with the selection battery, it is clear that some dimensions are not being measured. These include, affective characteristics (like emotional stability, assertiveness and sense of responsibility), psycho-social characteristics (such as leadership, team work and discipline) and primary management ability (for example, decision-making and strategic planning). There appears to be an emphasis on cognitive abilities and perceptual-motor skills.

It would probably make sense to eliminate those instruments that did not correlate or correlated poorly with the Ground School Phase results. However, it was shown that the Blox test, for example, when included in the total battery increases the correlation with the Ground School Phase results. The decision to remove an

instrument from the total battery should not entirely rest on statistical evidence, but should also have logical and theoretical merit (Tabachnick & Fidell, 1996).

The Blox test's negative beta value is most probably a "suppressor variable" (Tabachnick & Fidell, 1996), characterised by a high correlation with one of the other predictors and a lack of association with the criterion.

Since phases 1 and 2 of the Vienna Determination test form an integral and inseparable part of the entire test, it would be impossible to remove these phases from the Vienna test. One could however, in the light of the statistics, attach greater value to the Vienna Phase 3 during selection, due to its better correlation with the Ground School Phase results.

A comprehensive predictive validation study should be undertaken that would include the whole selection battery and hopefully a bigger sample, before any major policy decisions are taken concerning the current selection battery in general and pilot selection in particular. A validation study that includes the practical training results of trainee pilots, could also be explored.

In the light of the results of the study, it is essential to monitor whether the selection procedure has produced the ideal candidates, as anticipated (Jordaan & Jordaan, 1986). Hence, the importance of evaluating the selection procedure on a regular basis. Gerber et al., (1998) propose several areas on which to focus, in order to evaluate and possibly improve selection procedures. These areas include the following:

- ◆ Has a well-defined selection policy been developed?
 - ◆ Why are the current employment standards being used? How are they related to actual performance in the job?
 - ◆ Are accurate records being kept of the reasons why each candidate has been rejected?
 - ◆ What percentage of applicants has been employed?
-

- ◆ What contribution does each step in the selection process make towards the entire programme?
- ◆ How much does each of the steps in the selection process cost?
- ◆ Has every selection tool been properly validated?
- ◆ What percentage of the newly appointed employees is dismissed during the trial period?
- ◆ Can the selection process be successfully defended in court?
- ◆ Is there a correlation between the degree of success in the job and the predictions made during selection?
- ◆ Is there an exit interview to determine how well employees and jobs were matched?

By evaluating their performance in each of these areas, the South African Air Force will be in a position to ensure that the right sort and quality of candidates are attracted and appointed as pilots in the organisation.

In the light of these final conclusions, it is clear that the aims of this study in respect of the literature review and the empirical study have been addressed.

In respect to the aims of the literature review:

- ◆ The nature and value of scientific selection have been defined and discussed (see paragraph 2.1).
 - ◆ The steps in developing a scientific selection battery have been presented (see paragraph 2.2).
 - ◆ The concept of validity and the procedure of validating a selection battery have been discussed (see paragraph 2.3 and 2.4).
 - ◆ In the absence of an updated psychological profile of a pilot in the SAAF, the most recent one has been presented; the primary constructs being measured by the current selection battery have also been highlighted (see paragraph 2.5).
 - ◆ The dependent and independent variables of the study are discussed in chapter three.
-

In respect to the aim of the empirical study:

- ◆ The predictive validity of the current trainee pilot selection battery (in order to measure training success) has been determined (see chapter four).
- ◆ The limitations of the study will be discussed (see paragraph 5.2) and recommendations for future research will be presented (paragraph 5.3).

4.5 CONCLUSION

In this chapter, the results of the empirical study were reported and discussed. The results of the empirical study support the hypothesis that the total selection battery is valid for the prediction of training success of trainee pilots during the Ground School Phase.

In the final chapter, conclusions are made, limitations of the study are discussed and recommendations made for future research purposes.

CHAPTER 5

CONCLUSION, LIMITATIONS AND RECOMMENDATIONS

The aim of this chapter is twofold. Firstly, to highlight some of the conclusions that can be drawn from the results and discussion of the empirical study. Secondly, to discuss the limitations of the study and subsequently, make specific recommendations for future research purposes.

5.1 GENERAL CONCLUSIONS

The general aim of the study was to empirically and statistically determine the predictive validity of specific instruments (i.e. Blox, Advanced Ravens and Vienna Determination Test) which are part of the battery for the selection of trainee pilots in the South African Air Force.

On the basis of the results of the empirical study, the following conclusions were made:

- ◆ It was found that only the Advanced Ravens test and the Vienna test Phase 3 have a positive correlation with the Ground School Phase results of the trainee pilots.
- ◆ The total selection battery appears to be a good predictor of the Ground School Phase results.
- ◆ The Blox test and the Vienna Determination test phases 1 and 2 individually, do not appear to be good predictors of the Ground School Phase results.
- ◆ The respective beta weights of the individual instruments confirm that the Advanced Ravens test is the strongest predictor of Ground School Phase results, followed by the Blox test and the Vienna Determination test phase 3.

The value of the study was its confirmation that a positive relationship exists between some instruments of the trainee pilot selection battery and the trainees' Ground School Phase results. It also provided critical information that can be used to make policy decisions concerning trainee pilot selection.

5.2 LIMITATIONS OF THE STUDY

Limitations of the study in respect of the *literature review* can be delineated as follows:

- ◆ Confidential documents/research restricting the flow of information

A lot of research has been done on pilot selection in the military environment. Unfortunately, it was difficult, at times, to access this body of knowledge due to the confidential nature and sensitivity of some of the information.

- ◆ Reliability and validity of criteria

The reliability and validity of the candidates' final examination (which was used as the criterion variable in this study) has not been confirmed.

Limitations of the study in respect of the *empirical study* can be delineated as follows:

- ◆ Sample

The sample was too homogeneous in terms of culture group and gender, with the result that it was difficult to make assumptions and draw conclusions about other culture and gender groups e.g. women and black candidates.

- ◆ Data set

The cross-matching of the candidates and their results posed some challenges, since the candidates' Ground School Phase results (18 months later), was used as criterion for the empirical study.

- ◆ Pre-selection of candidates

Only those candidates who passed the psychometric evaluation (cut-off points determined by the Military Psychological Institute, based on specific norms for black and white candidates) were included for the Ground School Phase training. This pre-selection resulted in restriction of range, possibly lowering the correlation results (Kisser, 1986).

- ◆ Limited practical training as criterion

Due to the nature of their training, limited practical training results were used as criterion measure.

5.3 RECOMMENDATIONS FOR FUTURE RESEARCH

The following recommendations may be incorporated into future research:

- ◆ Database

A comprehensive, up-to-date database pertaining to pilot selection and training should be designed and maintained.

- ◆ Validation studies

Validation studies should be undertaken on a regular basis as the pilot population increases over the years.

- ◆ Validation study on whole battery

The results of this study should be followed up by another study that will include a predictive study on the whole selection battery (i.e. including the simulation exercises that are part of the total battery).

- ◆ Retain valid instruments and include new ones

Those instruments that show a positive correlation with the candidates' Ground School Phase results should be retained as part of a new trainee pilot selection battery, as well as the inclusion of new instruments. Of particular concern is the fact that the current battery does not measure the personality traits of prospective candidates.

- ◆ Sample size

The size of the sample should be increased both in terms of culture group and gender, where possible. One way of enlarging the sample is to promote pilot training.

- ◆ Promote pilot training

Actively promote pilot training among previously disadvantaged students, thereby, enlarging and enriching the applicant pool.

- ◆ Undertake a qualitative study

The possibility of undertaking a qualitative study should be explored in order to tap into the experiences of successful previously disadvantaged students in terms of the selection process, Ground School Phase and practical training.

- ◆ Normative comparisons

Normative comparisons can be done examining differences in scores obtained by various culture and gender groups, provided that adequate sample sizes can be obtained.

- ◆ Update job profile

The current job profile of a pilot in the South African Air Force should be updated.

5.4 CONCLUSION

This chapter provided a summary of the main findings of the empirical study, delineated the limitations the researcher encountered during the study and concluded by proposing specific recommendations for future research purposes.

REFERENCES

- Abrahams, N.M. & Alf, E.F. (1993). The treatment of failures in validation research. *Military Psychology*, 5(4), 235-249.
- Anastasi, A. (1988). *Psychological testing*. (6th ed). New York: MacMillan.
- Anastasi, A. & Urbina, S. (1997). *Psychological testing*. (7th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Aspeling, E.G. (1990). *Vlieënierskeuringstrategieë vir die 1990s*. Pretoria: National Institute for Personnel Research.
- Babbie, E. (1989). *The practice of social research*. (5th ed.). Belmont: Wadsworth.
- Bendix, S. (1996). *Industrial Psychology in the new South Africa*. Cape Town: Juta.
- Borack, J. (1994). Estimating predictive validity when range restriction due to selection and attrition is present. *Military Psychology*, 6(3), 193-204.
- Cascio, W.F. (1985). *Applied Psychology in Personnel Management*. New York: Reston Publishing Company.
- Cascio, W.F. (1991). *Applied Psychology in Personnel Management*. Englewood Cliffs: Prentice-Hall.
- Chmiel, N. (2000). *Introduction to Work and Organisational Psychology*. Cambridge: Blackwell.
- Cilliers, F.V.N. (1985). *Persoonlikheidspsigologie*. Potchefstroom: Wesvalia.

- Cronbach, L.J. (1994). *The essentials of psychological testing*. New York: Harper Collins.
- De la Rey, R.P. (1978). *Statistiese metodes in sielkundige navorsing*. University of Pretoria.
- Du Toit, H.J. (1989). *Die validering van 'n keuringsbattery vir hyskraandrywers*. Unpublished MA Thesis. University of Stellenbosch.
- Employment Equity Act, No 56 (1998). *Government Gazette*, 400 (19370).
- Forster, M.R. (2000). Key concepts in model selection: Performance and generalizability. *Journal of Mathematical Psychology*, 44(1), 205-231.
- Gatewood, R.D. & Field, H.S. (1987). *Human resource selection*. Chicago: Dryden Press.
- Gerber, P.D. , Nel, P.S. & Van Dyk, P.S. (1998). *Human resources management*. (4th ed.). Halfway House: International Thomson.
- Ghiselli, E.E. (1984). *Theory of psychological measurement*. New York: McGraw-Hill.
- Gordon, J. (1999). *Organisational behaviour: A diagnostic approach*. New York: Prentice-Hall.
- Gregory, R.J. (1996). *Psychological testing*. Boston: Allyn and Bacon.
- Herholdt, W. (1977). *'n Keuringsprogram vir Argitekstudente*. Unpublished Doctoral Thesis. University of Pretoria.
- Heunis, G. (1998). *Proposed changes to the process of pilot selection*. Cape Town: Langebaanweg.

- Hoffman, C.C. & McPhail, S.M. (1998). Exploring options for supporting test use in situations precluding local validation. *Personnel Psychology*, 51(4), 987-1003.
- Holburn, P.T. (1992). *Blox Test Administrator's Manual*. Pretoria: HSRC.
- Howell, D.C. (1995). *Fundamental statistics for the behavioral sciences*. (3rd ed.). Belmont: Duxbury Press.
- Huysamen, G.C. (1980). *Sielkundige Meting – 'n inleiding*. University of Pretoria.
- Jacobs, C.D., Haasbroek, J.B. & Theron, S.W. (1992). *Effektiewe Navorsing*. University of Pretoria.
- Jones, L.J. (1995). *Essentials of psychological testing*. (5th ed.). New York: Harper Collins.
- Jordaan, W.J. & Jordaan, J.J. (1986). *Man in context*. (3rd ed.). Johannesburg: McGraw-Hill.
- Kerlinger, F.N. (1986). *Foundations of behavioural research*. (3rd edition). New York: CBS.
- Kisser, L.M. (1986). *Research methods*. (3rd ed.). New York: Holt, Rinehart & Winston.
- Klinvex, B. (1999). *Organizational psychology*. New York: Macmillan.
- Labour Relations Act, No 66 (1995). *Government Gazette*, 366 (16861).
- Lewis, C. (1990). *Employee selection*. New York: Stanley Thores.
- Louw, D.A. (1986). *Inleiding tot die psigologie*. (2nd ed.). Cape Town: Lexicon.

- Louw, D.A. (1991). *Human development*. Pretoria: HAUM.
- Lundin, R.W. (1996). *Theories and systems of psychology*. (5th ed.). Lexington: Heath.
- Makgati, C.K.M. (2000). *Proposed changes to the process of pilot selection in the SAAF*. Pretoria: MPI.
- McCall, R.B. (1990). *Fundamental statistics for behavioural sciences*. (5th ed.). New York: Harcourt Brace Jovanovich.
- McGee, J.K. (1979). *Measurement, design and analysis*. Hillsdale, NJ: Erlbaum.
- Meiring, D. (1995). *Die ontwikkeling en validering van 'n keuringsbattery vir lede van die afdeling binnelandse stabiliteit van die Suid-Afrikaanse polisie*. Unpublished MA Thesis. Pretoria: University of South Africa.
- Meyer, W.F., Moore, C., & Viljoen, H.G. (1990). *Personology from individual to ecosystem*. Johannesburg: Lexicon.
- Milkovich, G.T. & Glueck, W.F. (1985). *Personnel/Human Resource Management : A diagnostic approach*. Los Angeles: Business Publications.
- Morgan, G. (1980). Paradigms, metaphors and puzzle-solving in organisational theory. *Administrative Science Quarterly*, 25(4), 605-622.
- Morgan, P.D. (1980). *Practical research: Planning and design*. (5th ed.). New York: Macmillan.
- Mouton, S. & Marais, H.C. (1990). *Basic concepts: Methodology of the social sciences*. Pretoria: HSRC.

- Muchinsky, P.M. (1993). *Psychology applied to work*. (4th ed.). Pacific Grove: Brooks/Cole.
- Owen, K. & Taljaard, J.J. (1996). *Handbook for the use of Psychological and Scholastic Tests of IPER and the NIPR*. Pretoria: HSRC.
- Plug, C., Meyer, W.F., Louw, D.A. & Gouws, L.A. (1986). *Psigologie-woordeboek*. (2nd ed.). Johannesburg: Lexicon.
- Prinsloo, C.H. (1992). *Manual for the use of the Sixteen Personality Factor Questionnaire, South African 1992 version*. Pretoria: HSRC.
- Raven, J.C. & Court, J.H. (1985). *Manual for Ravens Progressive Matrices*. London: H.K. Lewis.
- Reber, A.S. (1988). *The penguin dictionary of psychology*. London: Penguin.
- Ree, M.J. (1994). Predicting job performance: Not much more than the g. *Journal of Applied Psychology*, 79(4), 518-524.
- Reitz, H.J. (1981). *Behaviour in organisations*. London: Unwin Hyman.
- Robbins, S.P. (1998). *Organizational behavior: Concepts, controversies and applications*. New Jersey: Prentice-Hall.
- Rosnow, R.L. & Rosenthal, R. (1996). *Beginning behavioral research: A conceptual primer* (2nd ed.). Englewood Cliffs: Prentice-Hall.
- Saunders, E. (2000). *Assessing human competence*. Randburg: Knowledge Resources.

- Scheffler, M. (1991). *Die ontwikkeling van beroepsprofiel vir voorligting aan en keuring van hoëvlakwerknemers*. Unpublished Doctoral Thesis. University of Pretoria.
- Schuhfried, G. (1997). *Vienna Determination Test Manual*. Vienna: DPSE.
- Selltiz, C., Jahoda, M., Deutsch, M. & Cook, S.W. (1965). *Research methods in social relations*. New York: Holt, Rinehart & Winston.
- Smit, G.J. (1986). *Psigometrika*. Pretoria : HAUM.
- Smit, G.J. (1990). *Psigometrika, aspekte van toetsgebruik*. Pretoria: HAUM.
- Society of Industrial Psychology. (1998). *Code of practice for psychological assessment in the workplace*. Pretoria: Society of Industrial Psychology.
- SPSS Base 9.0 Applications guide (1999). Chicago: SPSS Inc.
- Tabachnick, B.G. & Fidell, L.S. (1996). *Using multivariate statistics*. (3rd ed.). New York: Harper Collins.
- Van der Walt, L. (1997). *Die voorspellingsgeldigheid van sielkundige maatstawwe in 'n finansiële instelling*. Unpublished MA Thesis. University of Pretoria.
- Van Wyk, M.W. (1993). *Fairness in Selection: Some present and future legal implications*. Paper delivered at Psychometric Conference, 14-18 June. Midrand: Society of Industrial Psychology.
- Verguts, T. & De Boeck, P. (1997). *Predicting performance on Raven's Progressive Matrices Test*. (Article posted on the World Wide Web). Belgian Psychological Society. Retrieved 17 April 2002 from the World Wide Web: <http://www.vlb.ac.be.bps/docs/activities/bps97/prog/bps97-p05.html>.

Vinchur, A.J. (1993). The prediction of predictability revisited. *Educational and Psychological Measurement*, 53(4), 1085-1094.

Walton, R.E. (1987). *Managing conflict*. (2nd ed.). Reading, M.A.: Addison-Wesley.

Wheeler, H.L. (1993). *The fairness of an engineering selection battery in the mining industry*. Unpublished MA Thesis. Pretoria: University of South Africa.

Zimmerman, C.H. (1957). *Effective Personnel Selection Procedures*. London: Staple Press.