THE PREDICTIVE VALIDITY OF A SELECTION BATTERY FOR UNIVERSITY BRIDGING STUDENTS IN A PUBLIC SECTOR ORGANISATION

Philippus Petrus Hermanus Alberts,

Hons BA (Psychology)

Mini-dissertation submitted in partial fulfilment of the requirements for the degree Magister Artium in the Department of Industrial and Personnel Psychology at the North-West University

Supervisor: Professor PE Scholtz

Potchefstroom

2007

COMMENTS

The reader should bear the following in mind:

- The editorial style as well as the references referred to in this mini-dissertation follow the format prescribed by the Publication Manual (4th edition) of the American Psychological Association (APA). This practice is in line with the policy of the Programme in Industrial Psychology of the North-West University to use the APA style in all scientific documents as from January 1999.

ACKNOWLEDGEMENTS

Herewith I would like to thank the following key individuals and organisations which assisted and contributed to the completion of this mini-dissertation:

- My Lord and Saviour, for guiding me and blessing me with the ability to complete this study.
- Prof PE Scholtz, my study leader, for his tremendous patience, guidance, encouragement and contribution to this study.
- Big thanks to Jannie Hartzenberg, colleague and research psychologist for his help and effort in preparing my statistical processing.
- The participants in the research project for collecting and capturing the data.
- The Youth Foundation, and all those who co-operated and set aside time to participate in this study.
- My colleagues, dear friends, family and especially my wife Amanda, who believed in me throughout this study, who listened to my struggles, supported me and helped me. I am incredibly grateful.
- Thank you to Cecilia van der Walt for the professional manner in which she conducted the language editing.

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Title: The Predictive Validity of a Selection Battery for University Bridging Students in a Public Sector Organisation

Key terms: Predictive validity, selection battery, validity, reliability, selection battery, selection

South Africa has faced tremendous changes over the past decade, which has had a huge impact on the working environment. Organisations are compelled to address the societal disparities between various cultural groups. However, previously disadvantaged groups have had to face inequalities of the education system in the past, such as a lack of qualified teachers (especially in the natural sciences), and poor educational books and facilities. This has often resulted in poor grade 12 results. Social responsibility and social investment programmes are an attempt to rectify these inequalities.

The objective of this research was to investigate the validity of the current selection battery of the Youth Foundation Training Programme (YFTP) in terms of academic performance of the students on the bridging programme. A correlational design was used in this research in order to investigate predictive validity whereby data on the assessment procedure was collected at about the time applicants were hired. The scores obtained from the Advanced Progressive Matrices (APM), which forms part of the Raven's Progressive Matrices as well as the indices of the Potential Index Battery (PIB) tests, acted as the independent variables, while the Matric results of the participants served as the criterion measure of the dependent variable. The data was analysed using the Statistical Package for Social Sciences (SPSS) software programme by means of correlations and regression analyses.

The results showed that although the current selection battery used for the bridging students does indeed have some value, it only appears to be a poor predictor of the Matric results. Individually, the SpEEx tests used in the battery evidently were not good predictors of the Matric results, while the respective beta weights of the individual instruments did confirm that the APM was the strongest predictor.

Limitations were identified and recommendations for further research were discussed.

OPSOMMING

Titel: Die Voorspellingsgeldigheid van 'n Keuringsbattery in 'n Openbaresektor-organisasie

Sleutelterme: Voorspellingsgeldigheid, geldigheid, betroubaarheid, keuringsbattery, keuring

Oor die afgelope dekade het Suid-Afrika ongelooflik groot veranderinge ondergaan wat die werksomgewing in die land beïnvloed het. Organisasies is genoodsaak om die sosiale ongelykheid tussen verskillende kultuurgroepe onder die loep te neem. In die verlede het die voorheen benadeeldes aan die kortste ent getrek met ongelykhede in die onderwysstelsel, soos 'n tekort aan gekwalifiseerde onderwysers (veral in die natuurwetenskappe), fasiliteite en boeke. Dit het dikwels gelei tot swakker gehalte graad 12-leerlinge. Sosiale verantwoordelikheids- en investeringsprogramme is 'n poging om ongelykhede reg te stel.

Die doelstelling van hierdie navorsing was om die voorspellingsgeldigheid van die huidige keuringsbattery van die "Youth Foundation Training Programme" (YFTP) ooreenkomstig die akademiese prestasie van die studente in die oorbruggingsprogram te bepaal. 'n Korrelasie-ontwerp is in hierdie navorsing gebruik om ondersoek in te stel na voorspellingsgeldigheid, waardeur data oor die keuringsprosedures tydens of in die omgewing van die tyd dat applikante aangestel word, ingesamel is. Die toetspunte wat uit die Advanced Progressive Matrices (APM) behaal is, wat deel uitmaak van die Raven's Progressive Matrices, asook die indekse van die Potential Index Battery (PIB)-toetse was die onafhanklike veranderlikes, terwyl die Matriekresultate van die deelnemers as die kriteriummeting vir die afhanklike veranderlike gedien het. Die data is deur die 'Statistical Package for Social Science' (SPSS)-sagtewareprogram ontleed deur van korrelasies en regressie analises gebruik te maak.

Die resultate toon dat, alhoewel die huidige keuringsbattery wat vir die oorbruggingstudente gebruik word, inderdaad bepaalde waarde inhou, dit slegs 'n swak voorspeller van Matriekresultate blyk te wees. Geen indekse van die SpEEx-toets wat in die battery gebruik is, het duidelik individueel matriekresultate betekenisvol voorspel nie, terwyl die beta-gewigte van die individuele instrumente wel bevestig het dat die APM die sterkste voorspeller is.

Tekortkominge is geïdentifiseer, en aanbevelings is gemaak vir toekomstige navorsing.

CHAPTER ONE

PROBLEM STATEMENT, AIMS AND OUTLINE OF THE RESEARCH

INTRODUCTION

This research is concerned with the validation of selected indices of the Potential Index Batteries (PIB) and the Raven's Advanced Progressive Matrices (APM) as predictors of academic success. In this chapter the reader is orientated towards the problem statement, the research aims and the method of research.

1.1 PROBLEM STATEMENT

The world is moving towards accelerated change and escalating diversity in all spheres of life. In South Africa (SA), democratisation has created a new socio-political order. South Africa was relieved from isolation and once again became part of the global community of nations. External changes had an impact and will continue to impact on South Africa as part of the world, which is progressively becoming a 'global village'.

The relatively young socio-political order in South Africa has inherited old societal disparities between the various cultural groups. Organisations are invariably also affected, and attempts have been made to address this state of affairs by launching so-called social responsibility and social investment programmes. The population groups on which this research will be based are from these previously disadvantaged cultural groups of SA. It is expected that the participants in this study will benefit from one of the above-mentioned programmes.

The first impetus which triggered the implementation of one such a programme in this public sector organisation was the government's Reconstruction and Development Programme (RDP). "The RDP is primarily aimed at realizing the full potential of everyone in the country and providing sufficient opportunities for all to become economically independent." (Erasmus & Minnaar, 1995, p. 34.)

In the light of the inequalities of the education system in the past, poor grade 12 results often resulted from systemic inefficiencies, such as a lack of qualified teachers (especially in the natural sciences) and poor educational facilities. It is generally accepted that opportunities were not equally distributed in Apartheid South Africa and also that skill competency is highly influenced by prior opportunity and learning. This programme therefore intends to give those previously disadvantaged youth the opportunity to enhance their Matriculation results in specific subjects, while undergoing teaching in a more conducive environment. The programme is a university bridging programme and is sponsored by the Government, with a private company providing the actual training. The focused curricula will be Mathematics, Physical Science and Biology. After one year of bridging training the successful candidates will be financially assisted in entering universities for further studies. These studies will be pursued in careers in which a critical shortage of personnel exists in the public sector organisation.

Thus another impetus for the implementation of the programme must be highlighted. Attaining a representative composition of the South African population in highly specialised occupations, such as pilots, navigators, engineers, technical personnel and professional medical personnel by merely marketing and recruiting candidates did not deliver adequate results. The available pool of high performing candidates is simply too small to satisfy the country's needs. In South Africa, as elsewhere, African students are underrepresented in natural science-, engineering- and technology-based programmes (DACST, 1996). High performing African grade 12 pupils are lured away by commerce and industry with attractive financial promises upon completion of their school careers. Others obtain bursaries. Additional measures were necessary to representatively fill the highly specialised occupational posts. University bridging training and contractual obligations after qualifying at tertiary institutions was believed to be the best way to sustainably ensure the quality and quantity of personnel needed.

The targeted profile of prospective students was:

- Previously disadvantaged (Africans, Indians and Coloureds).
- Aged 17-24 years.
- Out of school for a maximum of 3 years.

- In possession of a valid Matric certificate with Mathematics and Physical Science as subjects.

From an applicant pool of approximately two thousand two hundred candidates, approximately one hundred and seventy students need to be selected. This places selection by means of psychometric tests in focus. It is hence in this field of application in psychology (psychometrics) where the problem manifests itself, and where the researcher aims at making his contribution.

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

A number of legal sources exist which govern the conduct of psychometric assessment in South Africa. These are: The Health Professions Act; the Constitution of the Republic of South Africa, 1996 (108 of 1996); the Labour Relations Act, (66 of 1995); and the Employment Equity Act (56 of 1998). The latter Act is especially relevant, as it states that "Psychological testing and other similar assessments of any employee are prohibited unless the test or assessment being used has been scientifically shown to be valid; reliable; can be applied fairly to all employees; and is not biased against any employee or group" (Employment Equity Act (56 of 1998, p. 16). Good scientific and professional practices would require that similar considerations be made when conducting psychometric assessment. The '*Principles for the Validation and Use of Personnel Selection Procedures*' (SIOP, 2003) provides principles regarding the conduct of selection and validation research as well as the application and use of selection procedures. Both good science and legislation thus emphasise the importance of reliability and validity in terms of psychometric assessment.

This ethical and law imperative with respect to psychometric assessment, the financial implication for the organisation, as well as the fact that the selection battery for the target group had not yet been validated made it necessary to generate validation studies.

The overriding problem which this research aims to address can be stated as follows: How well do specific psychometric instruments, currently used as part of a selection battery, predict performance in Matriculation examinations for a group of university bridging students?

1.2 RESEARCH QUESTIONS

Considering the problem statement above, the following research questions can be formulated:

- How to undertake a scientific validation study. What is validity and especially predictive validity with respect to psychometric instruments?
- What does the Advanced Progressive Matrices (APM) Test entail?
- What do the Potential Index Battery Tests (PIB) entail?
- What is the relationship between scores obtained on psychometric instruments and grade 12 results?
- To what extent do the applicable instruments predict/declare the variance in Matric performance?
- Which instruments should form part of the selection battery for future use?
- What are the limitations of the study?

1.3 AIMS OF THE RESEARCH

1.3.1 General aim

The general aim of this research is to establish the validity of the current selection battery in terms of academic performance of the students in the University Bridging Programme.

1.3.2 Specific aims

The specific research aims of this study in terms of the literature review are:

- To define validity, in particular predictive validity, and to determine how to undertake a statistically based validation exercise.
- To describe what the APM test entails.
- To describe the applicable indices from the PIB.

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

- To describe the relationships between the scores on the applicable psychometric tests and the Matriculation examinations.
- To determine whether the instruments show predictive validity in this environment.
- To determine to what extent the applicable instruments predict/declare the variance in Matric performance.
- To discuss limitations of this study and to formulate recommendations based on the current study in order to improve the current selection battery.

1.4 THEORETICAL ASSUMPTIONS OF THE RESEARCH

Research is always conducted within the context of a specific paradigm (Mouton & Marais, 1994). The paradigm plays a critical role in terms of demarcating the boundaries of the research and to formulate specific points of departure for the research. The most applicable paradigms and meta-theoretical assumptions of this study are discussed with reference to the theoretical assumptions and the disciplinary relationship of the research.

This research is situated within the field of industrial psychology and its field of application with specific emphasis 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). The tasks of the individual psychologist include the study of organisations and organisational behaviour, personnel recruitment and selection, human resource management, the study of consumer behaviour, research as well as psychological testing.

Organisational Psychology as a sub-discipline deals with the individual dimensions of organisational behaviour, group and interpersonal process, organisational structure and organisational development (Du Toit, 1989).

Psychometric tests are objective standardised measurements of certain areas in human behaviour (Smit, 1996). Plug, Meyer, Louw and Gouws (1986) refer to psychometrics as the study of aspects of psychological measurement which focuses on the development and implementation of mathematical and statistical procedures.

According to Blake (1983), selection can be defined by the process of choosing, from those available, the person or people who best meet the requirements of a position or positions vacant within the organisation.

1.5 RESEARCH DESIGN

A research design arranges the conditions for collection and analysis of data in a way the research purpose and economic viability are combined (Selltiz, Jahoda, Deutsch, & Cook, 1976).

The quality of a good researcher is that he/she will always attempt to eliminate all those variables which might have an influence on the validity of the results. The research design fulfils a critical role in this regard. It helps to enhance the internal and external validity of the research findings (Mouton & Marais, 1994). The purpose of the research design is to determine whether the identified independent variables have an impact on the identified dependent variable (Huysamen, 1994).

The design used will be the concurrent design by means of which data on the assessment procedure is collected at or about the time applicants are hired (Society for Industrial Psychology, 1998). The scores obtained from the APM test as well as the indices of the PIB test (SpEEx 100, 200, 301, 302, 400, 1600) will act as independent variables and predictors. In order to determine predictive validity of any test, a valid criterion must be identified and made available in numeric format (Huysamen, 1994). The final year examination marks of the participants will serve as the criterion measure or the dependent variable. Marks obtained in the final year examination will be the most reliable, valid and objective indicator of performance as it is a national examination, which is externally compiled and marked. A combination average between two subjects (Physical Science and Mathematics) will be calculated and will serve as criteria on which this study will focus. The term, Matric examination results, which is used in this research, will thus refer to the above-mentioned combination average.

The research will take on a correlational format. A correlation will be drawn between the criterion and the predictor. The goal of correlational research is to determine the relationship between two variables and to determine whether the direction is positive or negative. Thus the main goal is to determine whether a correlation exists between the Matric examination results and the selection battery. Correlational research allows the researcher to simultaneously determine the degree and direction of a relationship with a single statistic (Kerlinger & Lee, 2000).

Regression, which will also be used in this research, is a technique which allows one to assess the relationship between one independent variable and several dependent variables (Tabachnick & Fidell, 1996). When using regression, one can correlate the independent variables with one another and with the dependent variables, which will make the research more experiential. The goal of regression is to arrive at the best set of regression coefficients for the independent variables which bring the y values predicted as closely as possible to the y values obtained by measurement.

The regression coefficients accomplish the following (Tabachnick & Fidell, 1996, p. 128):

- They minimize deviations between predicted and obtained y values.
- They optimise the correlation between the predicted and obtained y values.

The unit of analysis is the individual participant in the University Bridging Programme. The sample size is 173 people consisting of 136 males and 37 females. The sample used is equivalent to the entire population under investigation. Furthermore, the population consists of African, Asian and Coloured individuals of various ethnic origins and home languages from across South Africa.

In this research the literature review will be presented in a qualitative manner, and the empirical study will be presented in a quantitative, descriptive way.

1.6 SELECTION BATTERY

The selection battery which was administered to the participants consisted of two psychometric instruments:

- The *Advanced Progressive Matrices Test* (APM) provides a means of assessing more accurately a person's speed of intellectual work and is used for people over 11 years of age and of average or above average intellectual ability. 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). In order to assess a person's intellectual efficiency in the sense of his speed in producing accurate work, the test is administered within a specific time limit. Administering the test without time restriction would then give an indication of intellectual capacity (Raven, Raven & Court, 1998a). As this test is designed to differentiate between people of superior intellectual ability, it is often used to select staff for high-level technical or managerial positions.

  The APM consists of Set I and II. Set I consists of only 12 problems, which is followed immediately by set II, which 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 fast as possible (Raven & Court, 1985).

- The *Situation Specific Evaluation Expert* (SpEEx) from the Potential Index Battery (PIB) was developed in order to predict performance and success in the workplace (Erasmus & Minnaar, 1995). It has since further developed into a comprehensive organisational development system. The PIB is a series of culturally fair, computerised, flexible and comprehensive tests, aimed at illiterate, semi-literate and academically advanced individuals. The specific indices of the PIB used in this study include the SpEEx 100, 200, 301, 302, 400 and 1600. These indices from the PIB form part of the cognitive cluster of competencies. The other clusters of competencies are the social, emotional and conative competencies.

1.7 DATA ANALYSIS

The data will be analysed using the Statistical Package for Social Science (SPSS) software programme. Predictive validity will be determined for the purpose of this study. To achieve this, the data will be statistically analysed by means of correlations and regression analysis.

1.8 RESEARCH PROCEDURE

Phase 1 (Literature Review)

Step 1 Validity

Definitions, types of validity and the process for conducting a scientific validation study will be discussed.

Step 2 The Advanced Progressive Matrices Test.

The purpose, description and psychometric properties of this test will be discussed. A motivation for inclusion in the selection battery will be provided.

Step 3 The Potential Index Batteries.

The purpose, description and psychometric properties of the specific indices will be discussed. A motivation for inclusion in the selection battery will be provided.

Step 4 Bridging Training.

An overview of the purpose, curricula and examinations will be presented.

Phase 2 (Empirical Research)

Step 1 Description of the population and sample (see Chapter 1).

Step 2 Selection and motivation of the psychometric instruments (see Chapter 3).

Step 3 Data collection (see Chapter 3).

Step 4 Data analyses (see Chapter 4).

Step 5 Reporting and interpretation of the results (see Chapter 4).

Step 6 Conclusions, limitations and recommendations will be discussed (see Chapter 5).

1.9 CHAPTER DIVISION

The chapters of the study will be presented in the following sequence:

Chapter 2: Validation of psychometric tests.

Chapter 3: The selection test battery and the criteria will be discussed, as well as a description of the empirical study

Chapter 4: Empirical study with interpretation and discussion of the results.

Chapter 5: Conclusions, limitations and recommendations.

1.10 SUMMARY

In Chapter 1, the overall problem statement and research questions were set. This was followed by the aims and the theoretical assumptions of the research. The research design, the selection battery and method of analysis to be used were briefly discussed and concluded by the chapter division. The outline of the research is hereby concluded.

In Chapter 2, the process of scientifically validating a psychometric test will be clarified.

CHAPTER TWO

VALIDATION OF PSYCHOMETRIC TESTS

INTRODUCTION

This chapter will be concerned with the scientific validation of psychometric tests. The different types of validity, with emphasis on predictive validity, will be clarified. The statistical procedures of determining each kind will also be discussed.

2.1 TEST VALIDITY

In a broader sense, it is important to bear in mind that the selection board's decision is that which should be valid. It is thus the validity and fairness of the final selection decision which is of primary importance. In this regard the Standards for Educational and Psychological Testing, as published by the American Educational Research Association, American Psychological Association and the National Council on Measurement in Education (American Educational Research Association, American Psychological Association and the National Council on Measurement in Education, 1999, p. 9), discuss validity as follows:

> Validity is the most important consideration in test evaluation. The concept refers to the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores. Test validation is the process of accumulating evidence to support such inferences. A variety of inferences may be made from scores produced by a given test, and there are many ways of accumulating evidence to support any particular inference. Validity, however, is a unitary concept. Although evidence may be accumulated in many ways, validity always refers to the degree to which that evidence supports the inferences that are made from the scores. The inferences regarding specific uses of a test are validated, not the test itself.

Validity is complex, controversial and important in behavioural research (Kerlinger & Lee, 2000, p. 665). According to the *'Principles for the Validation and Use of Personnel Selection Procedures'* (SIOP, 2003), validity is seen as the most important consideration in developing and evaluating selection procedures.

However, in conventional usage and in a more narrow meaning of the term, the validity of a psychometric test gives an indication of whether the test itself measures what it is supposed to measure. The validity of a psychometric test results from and is dependent on the psychometric properties of the specific test.

According to Kerlinger and Lee (2000), there is no single validity form, and a test or scale is only valid for the scientific or practical purpose of its user. Thus, depending on the specific use of the test, different types of validity exist. Each of these types of validity has a different meaning and use. It can thus be summarised that a psychometric test must be validated in each specific situation where it is used and that the validity of a test is not an inherent, set or permanent quality. Two very important definitions of validity are the following:

- "The validity of a measuring instrument may be defined as the extent to which differences in scores on it may reflect true differences among individuals on the same characteristic that we seek to measure rather than constant random errors." (Selltiz et al., 1976, p. 168.)

- "Validity refers to the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration." (Babbie, 1992, p. 132.)

The following paragraphs will give a description of the definition, purpose, procedure and practical implications of each type of validity.

2.1.1 Content Validity

Babbie (1992, p. 133) defines content validity as follows: "Content validity refers to the degree to which a measure covers the range of meanings included within a concept."

It can thus be described as a method to determine whether test or scale items represent the behavioural aspect which they are supposed to measure.

The evaluation of the content validity of a psychometric instrument implies the logical investigation of the test content as well as the methods used in constructing the instrument. It is more appropriate to investigate the content validity of an instrument before construction is completed.

The following steps are essential in determining the content validity of a test:

1. The relevant universe of items must be defined in terms of tasks and situations with which the subject (testee) may be confronted.

2. The total universe of items must be systematically divided into subdivisions.

3. A probable sample of tasks or situations for each category must be assembled.

4. The selected tasks or situations must be written as questions.

Schaap (1997) points out that an instrument is considered biased in content when it demonstrates to be relatively more difficult for members of one group, when both groups have a similar measure of the underlying ability, and no reasonable theoretical rationale exists to explain group differences on the item or scale in question.

2.1.2 Criterion-related Validity

Criterion-related validity stands in relation to the use of measuring instruments which are used to make practical decisions concerning aspects such as the selection of applicants for positions. "Criterion-related validity is studied by comparing test or scale scores with one or more external variables, or criteria, known or believed to measure the attribute under study." (Kerlinger & Lee, 2000, p. 668.) This type of validity can be divided into two types, namely predictive validity and concurrent validity. Both types are based on the same principle, namely the comparison of test data with independent criterion data.

2.1.3 Predictive Validity

Predictive validity refers to the accuracy with which a test or instrument enables you to predict some future behaviour or status of individuals (Huysamen, 1983). According to Walsh and Betz (1995), it indicates whether and how present performance on the test predicts future success on the criterion variable.

The purpose of this type of validity is summarised in its definition. An example of where it is of utmost importance, is in determining the effectiveness of selection batteries for the prediction of success in training or job performance. However, it is imperative that the measures of the criteria such as success during training or job performance are valid measures of the criteria themselves.

In order to determine the predictive validity of any test/instrument, the following steps serve as a guideline (Meiring, 1995, p. 11):

1. A valid criterion (e.g. measurement of behaviour or individual's status) must be identified and made available in numeric format.

2. Each individual's performance on the psychometric instruments must be linked to their performance or rating on the criterion, for example by means of their identity number or surname and initials.

3. The test/instrument must be administered to a relatively large sample (100 depending on the number of tests or instruments included in the battery), which is representative of the population on which the test/instrument is to be used.

4. It is, however, important to realise that the results of such a study may only be generalised to people and criteria which correspond to those used in the validity study.

5. A statistical comparison is done between test scores and the criterion score, which serves to represent success.

A distinction is made between the true or conceptual criterion, and the available or operational measure of the criterion. The conceptual criterion refers to some standard or another in terms of which an individual's behaviour must be evaluated as successful or unsuccessful (Huysamen, 1983). This standard is often not directly measurable. An indirect measure of career success

could for instance be salary. It is clear that this measure can only be an indication of the criterion and not the criterion itself. A problem often experienced is finding a suitable measure of the criterion. It is important that the criterion be reliable and valid.
When an objective measure of the criterion behaviour is lacking, ratings of individuals' behaviour are often used. In such cases it is important to try to limit the personal influence of the raters. Usually more than one observer and multiple evaluation scales are used in an attempt to limit the problem. Researchers most commonly experience criterion contamination in the workplace. Brown (1983, p. 101) describes criterion contamination as follows: ".... the situation in which a person's criterion score is influenced by the rater's knowledge of his predictor score". This problem can lead to an artificial increase in the validity coefficient.

The most general methods of determining predictive validity of a test are:

- Validity coefficients

  Gregory (1996) and Rudner (1994) state that the most popular method of determining predictive validity of a test is by correlating test scores with criterion scores. The correlation is known as a validity coefficient. It must be noted that there is, however, no general indication of how high this coefficient should be (Smit, 1996).

- Contrast groups

  This method can be seen as an investigation into whether the test scores differentiate between contrasting groups divided on the basis of the criterion (Anastasi, 1988). A high and low performance group is selected on the basis of their scores on the criterion. It is then determined whether a significant difference exists between the two groups.

Bias in predictive or criterion related validity could have detrimental effects on both the individual and the organisation. This is due to the fact that the prediction of future performance behaviour is the most important purpose of assessment instruments. This prediction is a crucial consideration during selection decisions. Instruments used for selection purposes should predict future performance equally well for persons from the different racial components of society. Schaap (1997, p. 39) states in this regard: "Constant error in prediction as a function of membership from a particular group constitutes instrument bias. An instrument is unbiased if the results for all relevant sub-populations cluster equally well around a single regression line."

Finally, caution should be taken when interpreting coefficients from concurrent or predictive validity studies. The correlation is expressed in terms of a coefficient, which means that, if significant, a positive relationship or association can be established between the predictor and the criterion. This relationship is not a causal relationship. The implication for the current research will thus be that the research wishes to determine how accurately achievement in the mentioned Matric can be predicted as a consequence of achievement in the psychometric selection battery.

2.1.4 Concurrent Validity

"Concurrent validity can be described as a form of empirical validity which is determined by correlating test scores with criterion scores obtainable at the same time." (Plug et al., 1986, p. 266). It refers to the accuracy with which a test can identify or diagnose the current status of an individual's behaviour. Concurrent validity differs from predictive validity in the sense that the criterion data and the test data are available simultaneously.

This type of validity plays an important role when new instruments are developed which measure the same constructs measured by other older, reliable and valid instruments.

In order to determine concurrent validity, new and old tests measuring the same concepts should be administered simultaneously to a large representative sample of the population for which the tests are developed. The correlations between the scores on the new and old test are then analysed in order to determine to what extent the two tests measure the same construct. It must also be noted here that a high concurrent validity coefficient does not necessarily mean that the new test measures what we believe it to measure (Kerlinger & Lee, 2000). It could merely mean that the two tests in question cover the same theoretical framework or area.

Another procedure used to determine the concurrent validity entails determining how well a test distinguishes between individuals who are known to differ on a specific criterion.

2.1.5 Construct Validity

A construct is an imperceptible, hypothetical variable which forms part of a theory, developed to explain observable behaviour (Gregory, 1996). Almost all psychological concepts, for example intelligence, interest, attitude and performance motivation are hypothetical constructs. These constructs must be measured or quantified before any assumption (hypothesis) concerning relationships between these constructs can be tested. Construct validity is based on the way a measure relates to other variables within a system of theoretical relationships (Babbie, 1992).

In simple terms, construct validity can be defined as the extent to which a test measures the theoretical construct it is supposed to measure. Dane (1990, p. 259) defines it as follows: "Construct validity involves determining the extent to which a measure represents concepts it should represent and does not represent concepts it should not represent."

Construct validity is important when a test is developed (or an existing test is evaluated) for the purpose of investigating certain attributes or constructs which vary between individuals. An example of this would be the validation of a psychometric test. Research on a psychometric test would be aimed at determining whether it measures the construct it claims to measure, when tested on a sample. The same research could also be aimed at determining whether the test functions effectively amongst different cultural groups.

When an instrument is shown to measure different constructs for one cultural group than another or to measure the same construct but to a different degree of accuracy, bias is constituted. A non-biased instrument will reveal a high degree of similarity on its factorial structure across different cultural groups. There will also be similarity on the rank order of item difficulty within the instrument (Schaap, 1997). For instance, if one item is disproportionately more difficult for one cultural group compared to the difficulty of the same item by another cultural grouping, this item is biased or considered to be culturally loaded.

Construct validity cannot be determined by means of one single numerical index. A wide variety of methods are used to determine construct validity. These methods can be divided into two categories, namely intra-test methods and inter-test methods. A brief description of each will be given.

Intra-test methods: These methods are aimed at the investigation of the internal structure of the test (Smit, 1996). In other words the researcher looks at the expected pattern of responses, the internal structure of the instrument as well as the relationship between items or subscales of the instrument. These methods give the researcher information concerning the area of behaviour measured by the instrument and is usually obtained by means of factor analysis. This method, however, provides no information regarding the relationship between this construct and other variables. It is also possible to investigate whether an instrument measures the same construct in different groups, for instance in different cultural groups.

Inter-test methods: These methods imply the evaluation of the inter-correlations of several tests simultaneously. They are aimed at identifying commonalities and determining whether tests measure the same construct (Smit, 1996). These tests have to be administered simultaneously with the newly developed instrument. Two inter-test methods can be distinguished:

- Method of congruent/convergent validity: According to Dane (1990, p. 259) the definition of this type of validity is: "...the extent to which a measure correlates with existing measures of the same concept." The newly developed test is thus correlated with the existing test. High correlations give an indication that the two instruments measure the same construct.

- Method of discriminant validity: According to this viewpoint, a test is not only invalid when it does not correlate well with a test which measures the same construct, but is also invalid when it correlates too highly with a measure from which it is supposed to differ (Smit, 1996). For example, if a specific ability is supposed to differ between groups, t-tests or one-way analysis of variance can be used to confirm the construct validity of a test.

2.1.6 Face Validity

According to Dane (1990, p. 257), face validity refers to "consensus that a measure represents a particular concept". This kind of validity is usually not statistically calculated but rather based on the opinion of experts that the face value of the test or instrument appears as if it will be measuring what it is supposed to measure. Therefore it is also called validation by consensus.

2.2 THE PURPOSE OF VALIDATION

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 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 whom 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 which 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 successfully selected candidates (based on the criterion) can be increased by selecting a smaller, but according to the predictor, more promising group of candidates.

2.4 FACTORS INFLUENCING THE VALIDITY COEFFICIENT

As in the case of a correlation coefficient, the validity coefficient is influenced by any factor which influences the size of the correlation coefficient. These factors can be summarised as follows (Smit, 1996):

- **The range of the distribution of individual differences in the performance of the standardisation sample**

  Occasionally it is not possible for the researcher to determine the validity of the total range of performance of a specific test. As in the case of the reliability coefficient, the range of the distribution of individual differences also influences the validity coefficient. The more limited the range, the lower the validity coefficient (Smit, 1996).

- **The influence of test length on validity**

  An increase in the length of a test leads to an increase in the reliability of a test (Smit, 1996). There is a proportional relationship between the reliability and validity of a test, which implies that the lengthening of a test also increases the validity of the test.

- **The influence of the reliability of a test on the validity**

  It is generally accepted that if all factors are constant, the validity of a test is directly proportional to the reliability of a test (Smit, 1996). Ghiselli (1964, p. 353) states the following in this regard: "...we can see that as the reliability of either the predictor or the criterion becomes lower and lower, the validity becomes lower and lower..." Hence we can see that reliability limits validity and that, for optimal prediction, both the predictor and the criterion should be measured with as high a reliability as possible. Helmstadtler (1964, p. 85) explains the relationship between reliability and validity as follows: "The maximum possible validity (in this case, between a test and some independent measure of performance), is the square root of the reliability."

- **The influence of group heterogeneity on the validity of a test**

  Since the validity coefficient is a correlation coefficient, it is influenced by the heterogeneity of the group it is determined for. It is usually the case that the group, of which the criterion data is available, is selected according to certain variables. In as far as these variables are related to the predictor or criterion variables, this selection leads to a more homogeneous group than the original group, which is applicable to the criterion-related validity (Smit, 1996). This selection thus leads to a decrease in the validity coefficient. The reliability coefficient as well as the validity coefficient is influenced by the selection of a homogeneous subgroup of a population.

2.5 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 legislations include the Labour Relations (66 of 1995) and the Employment Equity Act (56 of 1998).

The Labour Relations Act (66 of 1995, p. 13) states that the overall purpose of the act is "the advancement of economic development, social justice, labour peace and the democratisation of the workplace". 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.

Organisations need to ensure that they comply with the Labour Relations Act (66 of 1995) 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 rests on the employer to establish that a specific practice is fair. The Employment Equity Act (56 of 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.

Employers and organisations should be sensitive about issues regarding bias and fairness, given South Africa's diverse multicultural 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 which 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.

2.6 THE PROCEDURE OF VALIDATING A SELECTION BATTERY

In selecting or promoting employees, the practitioner needs to answer a basic question: Do candidates, who perform better in 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. According to Klinvex (1999), the major steps in conducting a criterion-related validation study are job analysis, development of criterion measures of job performance, selection of predictors, composition of study sample, statistical analysis and the implementation of validity study results. Each of these steps will subsequently be discussed.

2.6.1 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 which will serve as the criteria of job performance to be "predicted" by the tests and to identify the appropriate selection instruments which will make up the trial test battery.

2.6.2 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 relevant to the organisation's core business and which validated tests seek to predict. Typical criterion measures are production data, personnel data and supervisory evaluations.

2.6.3 Selection of predictors

The term "predictor" refers to the selection instrument which 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. When selecting predictors it is important to keep the goal, which is prediction, in mind. According to Dane (1990, p. 7) prediction refers to: "...identifying relationships that enable us to speculate about one thing by knowing about some other thing." Examples of predictors are skill or ability tests, personality or interest inventories, knowledge tests, interviews and reference checks.

2.6.4 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. Klinvex (1999) further maintains that two conditions may render a criterion-related study technically infeasible. One constraint concerns 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 and 100 or more). "The larger the sample size, the smaller the standard error of the mean. And the smaller the standard error of the mean, the smaller the confidence interval about any estimate. To get a smaller confidence interval, select a larger sample." (Dane, 1990, p. 295). Thus, adequate variability in predictor and criterion scores as well as adequate sample size are threshold requirements for a criterion-related validation study.

2.6.5 Statistical analysis

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

- To summarise the data for ease of understanding. The relationship between test scores and criterion scores is expressed by the correlation coefficient.

- To "infer" by evaluating whether obtained results are statistically significant or whether they can be attributed to chance.

- 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.

2.6.6 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. Earlier, it was indicated that a thorough job analysis is suitable for not only test validation purposes, but also training and job evaluation. 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).

2.7. SUMMARY

In this chapter, the different types or aspects of psychometric test validity were discussed. The statistical methods used in order to determine each kind of validity as well as possible indications of bias concerning each type of validity were presented. Although, the types of validity are conceptually independent, they are considered to be practically interdependent (Gregory, 1996). It is imperative that scientific selection be conducted in accordance with our stringent labour legislation. 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. This emphasises the importance of this research. These specific research aims, namely to define validity, and in particular predictive validity, and to determine how to undertake a scientific validation exercise are hereby achieved.

In Chapter 3, the selection battery, which will serve as the predictor (independent variable) will be discussed. The criterion (dependent variable) is the results in the final Matriculation examination. These examinations were written at the year-end whilst undergoing foundation training throughout the year. The University bridging Bridging Programme will also be presented in Chapter 3.

CHAPTER THREE

EMPIRICAL STUDY

INTRODUCTION

In this chapter, the researcher will describe the predictor, i.e. the selection battery (APM and SpEEx indices of the PIB), as well as the criterion. The data collection procedure and a detailed discussion of the method employed for the empirical study will also be presented. The research hypotheses for this study will then be described.

3.1 STUDY POPULATION AND SAMPLE

A sample of convenience was used, since only information available on database was used for the empirical study. The population comprised bridging students who had already completed grade 12 and who had participated in the YFTP - a total of 173 students (100%).

The sample group consisted of previously disadvantaged students between ages 17 and 24 who had been out of school for a maximum of three years and who were in possession of a valid Matric certificate with Mathematics and Physical Science as subjects.

Table 3.1 depicts the study population. Here it is evident that the sample consisted mainly of African students (89,6%) with the majority of the participants being male (78,6%). Most of the participants were Setswana (23,1%) and isiZulu (22,5%)-speaking, with only 2.9% having English as their home language.

Table 3.1

Characteristics of the Participants

Item	Category	Frequency (Percentage)
Race	African	155 (89.6%)
	Asian	3 (1.7%)
	Coloured	15 (8.7%)
Gender	Male	136 (78.6%)
	Female	37 (21.4%)
Language	Afrikaans	13 (7.5%)
	English	5(2.9%)
	Ndebele	1 (0,6%)
	NSotho	3 (1,7%)
	Sepedi	2 (1,2%)
	Sesotho	8 (4,6%)
	Setswana	40 (23,1%)
	Siswati	31 (17,9%)
	Swazi	1 (0,6%)
	Tsonga	1 (0,6%)
	Venda	7 (4,0%)
	isiXhosa	22 (12,7%)
	isiZulu	39 (22,5%)

3.2 THE ADVANCED PROGRESSIVE MATRICES

3.2.1 Development and rationale

The Raven's Progressive Matrices (RPM) is a term used to describe both the Advanced Progressive matrices (APM) and Standard Progressive Matrices (SPM).

Raven et al. (1998a) describes the original purpose for developing the RPM and Vocabulary Tests for use in research on the genetic and environmental origins of mental defect. These tests sought to measure two components of the g-factor identified by Spearman, namely eductive and reproductive ability. He then defines eductive ability which is measured by the RPM as "... the ability to make meaning out of confusion; the ability to forge largely non-verbal constructs which make it easy to handle complexity." (Raven et al., 1998, p. 1). He defines reproductive ability as the ability which "... involves familiarity with a culture's store of explicit, largely verbal, information" (Raven et al., 1998, p. 1). The latter ability is measured by means of his Vocabulary Tests.

To summarise, the RPM taps "... something which might tentatively be called general conceptual ability" (Raven et al., 1998, p. 74).

3.2.2 Aim of the test

The APM was developed after the SPM as a mechanism to spread the scores of the more able. This was done after an increase in scores on the SPM over the years. The APM therefore gives an indication of higher-level eductive ability and assessing speed of accurate intellectual work (Raven et. al., 1998a). By imposing a time limit it can therefore be used to assess a person's 'intellectual efficiency' in the sense of his present speed of accurate intellectual work. This is generally, but always, closely related to his total capacity for orderly thinking. As a result, the two must not be confused with one another. Knowledge of a person's intellectual efficiency is particularly useful to assess a person's suitability for work in which quick, accurate judgements need to be made or when, in clinical work, a person's slowness of thinking has to be assessed (Raven & Court, 1985).

3.2.3 Description and administration

Set 1 consists of 12 problems only. It is used to provide the necessary training in the method of working. This is immediately followed by Set 2.

Set 2 consists of 36 problems, arranged in ascending order of difficulty. The series requires the examinee to choose which piece (from eight options) best completes a pattern series presented across three rows of designs. It is not necessary for everyone to attempt all problems before stopping. The time restriction for the completion of the test is 45 minutes (Raven & Court, 1985). The raw scores obtained by each student on the University Bridging Programme in the APM will serve as predictor in the empirical study (see Chapter 4).

3.2.4 Validity and reliability

As with the other versions of Raven's Progressive Matrices, the APM has been found to yield reliable scores as a measure of general intelligence, and it correlated 0,74 with the full-scale Wechsler Adult Intelligence Scale (WAIS) and 0,75 with the Otis I.Q. (McLaurin, Jenkins, Farrar, & Rumore, 1973). The internal consistency of the APM has been found to be substantial, with split-half reliabilities ranging from 0,8 to 0,9 (Alderton & Larson, 1990; Arthur & Day, 1994). The test-retest reliabilities have also been determined to be substantial ($r = .83$) (Bors & Stokes, 1998). The test manual reports a test-retest reliability of .91 for adults (Raven & Court, 1985).

According to a study on the construct validity of the Raven's APM for African and non-African engineering students in South Africa, the scores on the APM are as valid for Africans as they are for non-Africans (Rushton, 2004). For the African group, the mean $r=27$, $p<0,05$, and for the non-African group, the mean $r=27$, $p<0,05$. Although the intercepts of the regression lines for the two groups were significantly different, their slopes were not (Rushton, 2004).

3.2.5 Motivation for inclusion in battery

A brief reasoning on the selection of the APM is as follows:

- Firstly Raven et al. (1998, p. 5) states "General Intelligence and g have predictive validities of approximately .7 within the so-called 'academic' area". Since this research aims at validating a selection battery for students also in an academic environment and the test is included in almost all selection batteries in the current organisation, it was decided to put the APM on trial.

- Secondly the APM is regarded as being culture fair and relatively independent of language skills, making it appropriate for use on any cultural group. Since almost all the students participating in this research spoke English as their second language, the APM was ideally suited to diminish the cultural loading associated with verbal tests.

- Finally, but most decidedly, a job analysis was done by the researcher. The result was that three of the most important dimensions for success in the targeted position of a student were so-called ‘cognitive dimensions’.

3.3 THE POTENTIAL INDEX BATTERY (PIB)

3.3.1 Development of the PIB and brief description

Erasmus (2001, p. 2) states: “The new J P EXPERT and SpEEx and their older forebears, CSIP and PIB, are extremely comprehensive - most probably the most comprehensive one-stop HR systems of their kind...”

The PIB is a registered South African psychological test which was developed by Erasmus and Minnaar in 1995 for the purpose of establishing potential in areas of human performance (Erasmus & Minnaar, 1995). The PIB is a series of culturally fair, computerised, flexible and comprehensive tests, aimed at illiterate, semi-literate and academically advanced individuals. It is divided into two broad categories, namely three visual tests and three pen-and-paper tests, and comprises six separate batteries, each aimed at a specific population. Each separate battery is divided into a number of indices (Erasmus & Minnaar, 1995). The sixty-five indices are aimed at screening potential in various cognitive, emotional and social dimensions (Erasmus & Minnaar, 1995). These dimensions are also defined by Erasmus (2001) as basic competencies or units of potential, and he adds that the total field of human capacity is included in these 67 basic competencies.

3.3.2 The Situation-Specific Evaluation Expert (SpEEx)

The aim is to provide a comprehensive assessment package suitable for the assessment and development of human potential (in the workplace) in the South African context. The computerised generic norms are based on the South African population. The system can also develop norms on a situation specific basis, i.e. in the user’s environment for populations defined by the user. (Erasmus, 2001).

As certain tests from the SpEEx were used in this research, the definitions of the dimensions measured by these specific tests are given below. The raw scores obtained on the tests by each student in the University Bridging Programme will serve as the predictor in the empirical study (see Chapter 4).

- **SpEEx 100: Conceptualisation**
  The potential or capacity to reason in spatial terms; to see the relationship between parts; to complete the picture; to envisage the whole or end-result; to anticipate the outcome.

- **SpEEx 200: Memory**
  The potential or capacity to remember.

- **SpEEx 301: Basic Calculations**
  The potential or capacity to work with numbers or figures.

- **SpEEx 302: Advanced Calculations**
  The potential or capacity to work and deal with numbers and figures of advanced complexity.

- **SpEEx 400: Observance**
  The potential or capacity to pay attention: to understand; to be sharp or alert.

- **SpEEx 1600: Reading Comprehension**
  The potential or capacity to read and to understand what has been read, clearly and objectively.

3.3.3. Psychometric Properties

Schaap (2001a) conducted a research study investigating the psychometric properties of the SpEEx. A reliability analysis was performed. Results from this study indicated that the overall mean coefficient for the cognitive scales was calculated at 0,8. Six scales from this cognitive cluster of competencies had reliability coefficients of 0,79 and higher. Three scales ranged from 0,75 to 0,77 and four scales had reliability coefficients ranging from 0,72 to 0,75.

In a comparative analysis between African and White subgroups, the results indicated an average reliability of 0,78 for the African group and 0,66 for the White group. The difference in the average measurement error for the African and White groups is of practical significance. The cognitive scales measures are, on average, close to being equally accurate for both groups under consideration. It is important to note that the difference in the mean scores for the cognitive scales for the African and white groups reflect differences in the educational levels of the groups, largely due to historical inequalities. It seems evident that there is an absence of construct bias for these groups, as the average correlation between the transformed z-values (0,89) for the cognitive scales is high.

Erasmus (2001) clearly states that the PIB is the forebear of the SpEEx. Inclusion of research results obtained from PIB studies is therefore applicable and contributive when considering the various properties of the SpEEx. Valuable information regarding psychometric qualities of the SpEEx can thus be obtained in this manner. Schaap (1997) concludes that the predictive validity coefficient obtained from the PIB battery is well related to other research findings. Cascio (1991) states that according to meta-analysis procedures whereby different research findings were summarised, it was found that the average predictive validity coefficients for cognitive related instruments for managerial success ranged from 0,25 to 0,3. It needs to be emphasized that a predictive validity coefficient as low as 0,2 can justify inclusion in a selection programme (Anastasi, 1988).

According to results obtained from Schaap's (2001a), Basic Calculations reflected the highest reliability study (0,91), while Advanced Calculations measured the lowest reliability, namely 0,72.

Table 3.2

Reliability Statistics for SpEEx

SpEEx Stats	N	Mean	SD	A
100 Conceptualisation	1689	20.17	6.24	0.90
200 Memory	1177	12.41	3.65	0.75
301 Basic Calculations	1099	12.20	5.46	0.91
302 Advanced Calculations	1062	8.90	3.41	0.72
400 Observance	1427	14.14	3.45	0.76
1600 Reading Comprehension	264	6.69	4.50	0.85

Table 3.1 explicates the reliability coefficients of specific tests used in the current research study.

3.3.4 Motivation for inclusion in battery

- The Public Sector organisation is already using the PIB in various settings. Thus the instrument's validity for selection purposes and predicting success in an academic setting had to be determined.

- The job analysis for the position of a student was done according to the Job Profiling Inventory (JPI), which forms part of the PIB.

- A large number of the SpEEx tests are non-verbal and considered to be culture free.

- It is a South African test, standardised on a representative sample of the population.

- According to the test developers, the principle according to which the test was developed was that the ability to perform a simple task indicates the potential to master a related but more complicated task. The SpEEx claims to measure this basic competency and this "equals size, capacity, potential, ability and input" (Erasmus, 2001, p. 6). Functional competency in PIB terms is the output, acquired skill through previous exposure or experience (Erasmus, 2001).

- Schaap (2001) concludes that from the research evidence pertaining to the SpEEx, it is clear that the scales can be of valuable assistance in identifying potential for specific positions in the workplace.

3.4. THE UNIVERSITY BRIDGING PROGRAMME

3.4.1 Background

The information summarised in paragraphs 3.4.1 to 3.4.3 on the bridging programme's content was provided by H. Jordaan (personal communication, 10 February 2002) who is the specific project manager for this programme.

Normalisation of the racial imbalance in the current Public Sector Organisation's highly specialised occupations is required to comply with affirmative action and equal opportunity imperatives (see Chapter 1). Apart from recruitment efforts, additional measures were necessary to ensure filling of posts in careers such as pilots, navigators, engineers, air space controllers, occupational therapists, physiotherapists, human movement personnel, doctors and technical personnel.

In order to meet these challenges, the Public Sector Organisation is embarking on a process to implement a Foundation Training Programme (FTP) for previously disadvantaged youth. This programme will firstly meet the requirement for the enhancement of the educational qualifications, especially the grade 12 Mathematics, Physical Science and Biology symbols of previously disadvantaged applicants who apply for appointment in highly specialised technical and health occupations. Secondly, the programme will be a contribution by the Public Sector Organisation towards general youth empowerment.

The Bridging Programme provided by the private company is one such a programme which aims at optimally preparing selected Matriculated candidates from underprivileged communities for tertiary education. The Programme was established in 1994, driven by the long-term social investment strategy of the organisation. The latter notes the shortage of technically skilled people in certain underprivileged communities and states that the University Bridging Programme focuses on human resource development and, in particular, it supports, promotes and enhances technical training.

Technical performance stimulates the economy of South Africa and this company, as a high-technology engineering company, will always invest in training technically skilled manpower.

The mission of the programme is to contribute to the future success capacity of underprivileged students by exposing them to a learning environment of quality technical teaching and life skills development.

Personal potential development is enhanced by:

- Supplying quality education and service through total commitment.

- Creating an environment and opportunities conducive to human development.

- Facilitating appropriate tuition and support according to the academic and personal needs of the students.

- Exposing students to the world of technology and life skills development.

- Establishing the personal values of responsibility and continuous self-development.

3.4.2 Training

The staff consists of a number of part-time teachers who are all highly qualified educationists with experience in the presentation of the grades 11 and 12 syllabi on higher grade. These individuals are selected for their commitment to high educational values and standards.

Apart from the academic skills essential for tertiary education and the world of work, much emphasis is placed on the personal development and growth of students. To this end, the curriculum is further enhanced by additional subjects and programmes such as:

- *English* - with the emphasis on communication skills in a technical/commercial environment.

- *Life Skills* - to improve self-confidence; to provide career guidance, as well as essential 'job skills', e.g. financial basics.

- *Technology* - practical exposure to a technical environment.

- *Computer Literacy* – Word processing (setting up reports, internal memos, CVs, letters of application, etc); Spreadsheets (setting
- up personal budget, processing data from experiments, etc.); Drawing software (basic designs); and Internet (e-mail, searching the web for subject related material).
- *Entrepreneurship* – Focus is placed on new and innovative opportunities leading to the forming of a new business.

- *Advanced Mathematics* (Calculus and Vector Algebra) - for students obtaining above 60% for Mathematics during the first examination in March.

- *Technical Drawing* (Technical Group) - the N1 syllabus is followed, and students write this subject through the Centurion Technical College.

- *Introduction to Statistics* (Commercial Group).

- *Basic Mathematics* (grades 8-10 Mathematics revision).

- *Leadership Development Programme* - Consists of two aspects: (1) Student Leadership Positions: Responsibility for various aspects of the smooth running of the college is entrusted to the students. (For example, students manage various aspects of the College, such as the choir, computer room, tuck-shop, library, recreation, cleaning, first aid and safety, school magazine, kitchen, class administration and school notice boards); and (2) Student Management Team (SMT), which consists of 8 members, elected by the students, and who are trained to act as line managers in the college. Each SMT member has line management function over 9 other students, and some specific Student Leadership Position portfolios.

- *Environmental awareness* - Students are taken on a nature conservation camp where they are exposed to nature and interpersonal development activities (team building, leadership styles, etc).

- *Sport* - Participation in company sport days and utilisation of company sport facilities is encouraged.

3.4.3 Academic Curriculum

Mathematics and Physical Science/Accounting are taught with the emphasis on providing a solid background for tertiary education, and upgrading grade 12 external examination marks.

Mathematics and Physical Science are presented as separate modules (each 4,5 hours per week), thus making students aware of different specialisation fields. The Modules are:

- **Technical Group**
  Algebra
  Geometry and Trigonometry
  Chemistry
  Physics

- **Commercial Group**
  Algebra
  Geometry and Trigonometry
  Accounting

In the technical group, much emphasis is placed on practical work.

The students have all previously written a grade 12 examination (Matric) on the subjects before entering this programme. The Department of Education made arrangements for a national preliminary examination in September 2001. These papers were externally set by the Department of Education on the correct standard. An examination timetable was forwarded to the schools. All students countrywide wrote these papers simultaneously. A memorandum was sent to the schools according to which teachers could mark the papers.

The subjects in which these students enrolled were Mathematics and Physical Science, and some students also in Biology. All students enrolled on higher grade. It will be these grade 12 examination marks which will serve as criterion in the study.

3.5 DATA COLLECTION PROCEDURE

A description of how the data of both the independent and dependent variables was collected is presented below.

3.5.1 Data collection of independent variables

The psychometric tests (APM, SpEEx 100, SpEEx200, SpEEx 301, SpEEx 302, SpEEx 400, and SpEEx 1600) from the selection battery were administered and scored by Psychometrists and Psychologists under standard testing conditions. The tests were administered in paper and pencil format in a single testing session according to the administration procedures of each test. 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.

3.5.2 Data collection of dependent variable

Having completed the year of training, the university bridging students wrote various grade 12 subjects (Mathematics, Biology and Physical Science). This is referred to as 'Matric examination results' or 'Matric results' in this study. These results were used as the dependent variable.

3.6 STATISTICAL ANALYSIS

The data was processed by using the SPSS software programme (SPSS Base 14.0, 2005). The data was analysed with SPSS using the Pearson Product–Moment Correlations and the Multiple Linear Regression Analysis techniques.

The purpose of statistical analysis is to test the hypotheses (presented in Chapter 3) that the psychometric results of the students on the different psychometric tests which constitute the selection battery, correlate significantly with the same students' Matric results.

The following statistical techniques and procedures were used:

3.6.1 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 & Lee, 2000).

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.

According to Huysamen (1983), correlation coefficients are often used because most psychology practitioners are familiar with it (as it forms part of their training), and also because it is a convenient way of summarising a relationship into one general descriptive term. Huysamen (1983) also states that the value of the statistics 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 and 0,2 is regarded as poor, a correlation between 0,3 and 0,4 mediocre and correlation above 0,5 as strong (Howell, 1995).

The *"statistical significance"* of certain results can be determined by means of statistical inference. 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 one another, 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.

Hence three aspects are important for this analysis:

- Independent Variable correlations, which is the correlations of the independent variables with one another;

- Bivariate regression analysis, which is the unique relationship of independent variables with the dependent variable (individual R square values for each subtest); and

- The total relationship of independent variables with the dependent variable, whereby R square is given for all the independent variables together as a group (by means of multivariate regression analysis).

The results of the empirical study are reported in the above-mentioned format. The tables, which are given in Chapter 4, can be interpreted as follows:

- R represents the Pearson correlation between the independent variables (APM and the selected SpEEx tests) and dependent variable (Matric examination results).

- R square indicates the percentage of the variance in the dependent variable explained by the independent variable(s). R square is multiplied by 100 in order to obtain the percentage variance explained by each subtest.

- 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 the chance variations of some independent variables (subtests) "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 in 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 larger 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).

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

3.6.2 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).

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

- It minimizes errors in prediction by ensuring that predictors are combined to obtain the most efficient (highest) estimate of subsequent performance;

- 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;

- 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);

- The multiple regression approach makes it possible for high scores 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 pose a problem. One should check for critical minimum performance levels in specific predictors.

Both linear correlations between individual independent variables and the dependent variable were determined. The rationale behind utilising this technique is to determine whether there is a linear relationship between the independent variables (individually) with the dependent variable, and also because 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 though it had entered the regression after all other independent variables had entered. Each independent variable is evaluated in terms of what it adds to the prediction of the dependent variable, which 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 (such as 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). This emphasises the importance of this technique in this study.

3.7 HYPOTHESES

3.7.1 Basic Hypotheses

The APM and the specific indices of the PIB are valid assessment instruments in predicting success for final Matriculation examinations for university bridging students.

3.7.2. Research Hypotheses

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

H0: Raven's Advanced Progressive Matrices are not statistically significant for the prediction of success of the university bridging students in the Matriculation examinations

H1: Raven's Advanced Progressive Matrices are statistically significant for the prediction of success of the university bridging students in the Matriculation examinations

H0: The selected SpEEx indices of the PIB is not statistically significant for the prediction of the success of the university bridging students in the Matriculation examinations

H1: The selected SpEEx indices of the PIB is statistically significant for the prediction of the success of the university bridging students in the Matriculation examinations

H0: The total selection battery is not statistically significant for the prediction of success of the university bridging students in the Matriculation examinations

H1: The total selection battery is statistically significant for the prediction of training success of the university bridging students in the Matriculation examinations

3.8 SUMMARY

This chapter outlined the method of investigation which was used for the empirical study. The various tests of the selection battery, used as the independent variables in the study, were discussed at length. The statistical processing of the data was discussed in terms of the various techniques used to determine the predictive validity of each of the psychological tests. The chapter concluded by posing the research hypotheses in respect of the validity of each of the psychological tests used in the selection battery for the selection of university bridging training.

Chapter 4 will deal with the reporting of the data and the interpretation of the results from the empirical study.

CHAPTER 4

DISCUSSION OF RESULTS

INTRODUCTION

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 Matric examination results. The results are interpreted and discussed 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 one.

4.1 RESULTS OF INDEPENDENT VARIABLE CORRELATIONS

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

Therefore three aspects are important regarding this analysis:

- *IV correlations*: Correlations of the independent variables with one another.
- *Total relationship of IVs with DV*: R^2 for all the IVs together as a group (by means of multivariate regression analysis with 'Enter' method).
- *Unique relationship of IVs with DV*: Individual R^2 values for each subtest (bivariate regression analysis).

The research results and discussions will be presented in the above-mentioned format. The descriptive statistics of the APM, SP100, SP200, SP301, SP302, SP400 and the SP1600 tests are given in Table 4.1.

Table 4.1

Descriptive Statistics of the APM, SP100, SP200, SP301, SP302, SP400 and SP1600

Item	Mean	SD	Skewness	Kurtosis
APM	19.48	6.42	-0.59	0.54
SP100	21,66	3.42	-1,72*	3,2*
SP200	14,72	3.17	-0.24	0,45
SP301	17,87	2,01	-2,34*	10,76*
SP302	10,27	2,87	1,19*	5,02*
SP400	15.4	2,57	-0,41	-0,07
SP1600	9.35	3,44	0,02	-0,67

*High skewness and kurtosis

The data was not available at the item level: hence could not be analysed for reliability and factorial validity. Previous reliability and validity studies have been conducted on both the APM and the SpEEx tests (McLaurin, Jenkins, Farrar, & Rumore, 1973; Alderton & Larson, 1990; Arthur & Day, 1994; Bors & Stokes, 1998; Schaap, 2001a) and were discussed in Chapter 3. These previously mentioned studies displayed acceptable psychometric properties in terms of reliability and factorial validity.

It is evident from Table 4.1 that the SP100, SP301 and SP302 showed relatively high skewness and kurtosis, while the other tests displayed low skewness and kurtosis.

The correlations between the independent variables are reported in Table 4.2, and the correlations between the independent variables and dependent variable are reported in Table 4.3.

Table 4.2

Correlations between Independent Variables

Item	1	2	3	4	5	6
1. APM	.	.	.	.	.	.
2. SP100	0,50*+	.	.	.	.	.
3. SP200	0,19	-0,02	.		.	.
4. SP301	0,31*+	0,24*	0,37*+	.	.	.
5. SP302	0,29*	0,16	0,26*	0,06	.	.
6. SP400	0,51*++	0,34*+	0,05	0,12	0,32*+	.
7. SP1600	0,22*	0,08	0,25*	0,22*	0,21*	0,13

* Statistical significance: p<0,05

+ Practical significance: r>0,30 (medium effect)

++Practical significance: r>0,50 (large effect)

Tabachnick and Fidell (1996, p. 118) define multicollinearity and singularity as follows: "Multicollinearity and singularity are problems with a correlation matrix which 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".

The above-mentioned authors continue by saying that when variables are multicollinear or singular, they contain redundant information and they are not at all needed in the same analysis. The reason being that they cause both logical and statistical problems by inflating the size of error terms and therefore weakening an analysis. No indication of multicollinearity or singularity was found with the correlations between independent variables (Table 4.2).

Table 4.2 indicates statistically significant positive correlations (practically significant, large effect) between APM and both SP100 and SP400, as well as with SP301 (medium effect). The APM also shows a statistically significant relationship with the SP1600. The SP100 is significantly related to the SP400 (practically significant, medium effect) and statistically significantly related to the SP301. The SP200 is significantly and practically related only to the SP301 (medium effect), and statistically significantly related to the SP302 and the SP1600. The SP301 is statistically significantly related to the SP1600. The SP302 is statistically significantly related to the SP400 (practically significant, medium effect) and the SP1600.

Table 4.3

Pearson Correlations between Independent Variables and the Dependent Variable

	Matric Results
1. APM	0,37**+
2. SP100	0,23**
3. SP200	0,03
4. SP301	0,23*
5. SP302	0,13
6. SP400	0,25*
7. SP1600	0,14

* Statistical significance: p<0,05

**Statistical significance: p<0,01

\+ Practical significance: r>0,30 (medium effect)

++Practical significance: r>0,50 (large effect)

Table 4.3 indicates that the APM is the only test which reflects a statistically and practically significant positive correlation of a medium effect in terms of the Matric results ($p \leq 0,01$). The rest of the selection battery (SP100, SP200, SP301, SP302, SP400 and SP1600) reflected no practically significant relationships. However, the SP100, SP301 and SP400 are all statistically significant predictors of Matric results. Since these tests do not reach practical significance, the conclusion is drawn that they are not good predictors of Matriculation results during the bridging training. The observed significance level (the p-value) is used as the basis for deciding on whether to reject the null hypothesis. It is a probability that a statistical result as extreme as the one observed would occur if the null hypothesis were true. If the observed significance level is small enough, usually less than 0,01, or in the case of the Social Sciences 0,05, the null hypothesis is rejected.

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

- The correlation between the APM and the Matric results is practically significant (r=0,37; $p \leq 0,01$) with a medium practical effect.

- None of the SpEEx tests correlated both statistically and practically significantly with the

Matric results ($p<0,05$).

- The Sp100, SP301 and SP400 are the only SpEEx tests from the battery to correlate statistically (but not practically) significantly with the Matric results ($p\leq 0,05$).

The APM provides an indication of a person's intellectual efficiency (e.g. speed of accurate intellectual work), which is a critical skill required in the success of the Matriculation examination, hence the practically significant relationship between this test and the Matric results (dependent variable) of the students.

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

A summary of the total predictive value of the independent variables (APM, SP100, SP200, SP301, SP302, SP400 and SP1600) with the dependent variable (Matric results) is displayed in Table 4.4.

Table 4.4

Total Predictive Values between all Independent Variables with the Dependent Variable

	R	R^2	Adjusted R^2	Std. Error of the Estimate	F	P
Total Battery	0,37[a]	0,14	0,05	12.98	1,65	0,14

a. Dependent Variable: Matric Results

* Statistical significance: $p<0,05$

R represents the Pearson correlation between the independent variables (selection battery) and the dependent variable (Matric results). The total selection battery has no statistically significant linear relationship with the dependent variable (R=0,37; $p\geq 0,05$). The relationship between the independent variables and the dependent variable is therefore relatively poor.

R^2 is the variance in the dependent variable (Matric results) explained by the independent variables. In order to get an idea of the percentage variance explained by the overall selection battery, the R^2 can be multiplied by 100. The Adjusted R^2 is an adjustment due to the fact that a large number of independent variables are used in this research. It is possible that R^2 will become artificially high, simply because some the chance variations of independent variables explain small parts of the variance of the dependent. At the extreme, when there are as many independents as cases in the sample, R^2 will always be 1,0. When there are a great many independents, adjusted R^2 may be noticeably lower. The greater the number of independents, the more the researcher is expected to report the adjusted coefficient, hence the report on adjusted R^2 for this analysis. The total selection battery explains 5% (adjusted R^2) of the variance in the dependent variable. Thus only a small part of the variance of the dependent variable is explained by the independent variables.

An indication of the individual beta weights is given in Table 4.5. 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 in Matric results. By examining these weights one will notice that they reflect the relative strengths of each independent variable.

Table 4.5

Individual Beta Weights

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
(Constant)	11.69	17.37		0,67	0,50
APM	0,44	0,31	0.22	1.41	0,16
SP100	6.42	0,47	0,02	0,14	0.89
SP200	-6.90	0.54	-0,02	0.13	0,90
SP301	0,72	0,79	0,12	0.91	0.37
SP302	0.16	0.54	0,03	0,29	0,77
SP400	0.58	0.78	0.10	0.74	0.46
SP1600	0,19	0.45	0,05	0,43	0,67

a. Dependent Variable: Matric Results

* Statistical significance: p<0,05

**Statistical significance: p<0,01

Table 4.5 indicates that not one of the tests show significant power ($p \leq 0,05$) in terms of predictability. The beta weights clearly show that none of the tests have a statistically significant linear correlation with the dependent variable (Matric results). It is thus important to look at the correlation between the relationship of independent variables and the dependent variable.

4.3 RESULTS OF INDIVIDUAL RELATIONSHIPS OF INDEPENDENT VARIABLES AND THE DEPENDENT VARIABLE

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^2 value. These results are reported in Table 4.5.

Table 4.6

Correlation between Independent Variables and the Dependent Variable

	R	R^2	Beta	Significance Level
APM	0,37	0,14	0,13	1%
SP100	0.23	0,05	0,04	5%
SP200	0,03	0,00	-0,01	N/S
SP301	0,23	0,05	0,05	5%
SP302	0.13	0,02	0,01	N/S
SP400	0.25	0,06	0,05	5%
SP1600	0.14	0,02	0.01	N/S

N/S = Not significant

Table 4.6 can be considered a summary of the total relationship of the independent variables (selection battery) with the dependent variable (Matric results).

Clearly, there is no significant relationship between the SP200, SP302 and SP1600 with the Matric results (see Table 4.6). To determine the individual beta weights of the selected independent variables, the above-mentioned three SpEEx tests were omitted (see Table 4.7).

Table 4.7

Individual Beta Weights of Selected Independent Variables

Model	Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
(Constant)	13.09	15.57		0,84	0,40
APM	0,53	0.28	0,26	1.87	0,07
SP100	0,05	0,44	0,01	0,11	0,91
SP301	0,74	0,69	0,12	1.08	0,29
SP400	0.55	0,65	0,10	0.85	0.40

a. Dependent Variable: Matric Results

Table 4.7 indicates that not one of the selected tests have statistically significant predictive power ($p \leq 0{,}05$). The Ravens is the only test which comes close to having statistically significant predictive power ($p=0{,}07$) with regard tot Matric results.

4.4 INTEGRATION OF RESULTS

The overall results of this study seem to indicate that the current selection battery used for the bridging students has very little practical value. Some of the major conclusions which can be drawn from the results of the study are the following:

- The APM statistically and practically significantly correlated (large effect) positively with both SP100 and SP400, as well as with SP301 (medium effect). The APM also shows a statistically significant relationship with the SP1600. The SP100 is significantly related to the SP400 (practically significant, medium effect) and statistically significantly related to the SP301. The SP200 is significantly and practically related only to the SP301 (medium effect) and statistically significantly related to the SP302 and the SP1600. The SP301 is statistically significantly related to the SP1600. The SP302 is statistically significantly related to the SP400 (practically significant, medium effect) and the SP1600 (see Table 4.2). The correlation between these tests was expected due to the fact that these tests all measure aspects of cognitive potential.

- The total selection battery has no significant linear relationship with the dependent variable (R=0,37; p≥0,05) (see Table 4.4).

- The independent variables collectively explain approximately only 5% of the variance in the dependent variable (Adjusted R^2 = 0,05) (see Table 4.4).

- The APM is the test from the selection battery which correlates highest with the Matric results (r=0,37) (see Table 4.6).

- The APM, SP100, SP301, and SP400 are the only tests which show a statistically significant relationship (p≤0,05) with the Matric results during the analysis of the entire battery (see Table 4.6). When these selected tests were analysed, not one showed statistically significant predictive power (p≤0,05).

- The beta weights support the conclusion that the APM is the strongest predictor of the Matric results (see Tables 4.6 and 4.7).

The significance levels (p-values) provide an indication of whether or not the null hypothesis can be rejected. A significance level of less than 0,01 or 0,05 facilitates the rejection of the null hypothesis (Anastasi & Urbina, 1997; Huysamen, 1983; Rosnow & Rosenthal, 1996).

Based on the results of this study, the null hypothesis regarding the predictive power of the APM can be rejected. The null hypothesis regarding the predictive power of the total selection battery can also be rejected. The battery can be adapted to increase its predictive validity in the light of the statistical significance of the APM.

It would probably make sense to eliminate those instruments which did not correlate, or correlated poorly, with the Matric results. The decision to remove an instrument from the total battery should not rest on statistical evidence entirely, but should also have logical and theoretical merit (Tabachnick & Fidell, 1996).

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). This highlights the importance of evaluating the selection procedure on a regular basis. Gerber, Nel, & van Dyk (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?

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

4.5 SUMMARY

In this chapter, the results of the empirical study were reported and discussed. The results of this study reject the null hypothesis regarding the predictive power of the APM. Based on the results of this study, the null hypothesis regarding the predictive power of the total selection battery can also be rejected. The battery can be adapted to increase its predictive validity, in the light of the statistical significance of the APM.

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

CHAPTER 5

CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS

This chapter comprises conclusions regarding the literature review and the empirical study according to the specific objectives. The limitations of the research are discussed, followed by recommendations for the research problem in the organisation, and finally, suggestions are made for future research.

5.1 CONCLUSIONS

The objective of this research was to investigate the validity of the current selection battery of the YFTP in terms of academic performance of the students on the bridging programme. A correlational design was used in order to investigate predictive validity in this research, whereby data on the assessment procedure are collected at or about the time applicants are hired. The scores obtained from the APM and the indices of the PIB tests acted as the independent variables, while the participant's Matric results served as the criterion measure of the dependent variable. The data were analysed using the SPSS software programme by means of correlations and regression analyses.

The question as to how well specific psychometric instruments, used as part of the selection battery, predict performance in Matriculation examinations for a group of university bridging students was the main focus of this study. The general aim therefore was to establish the validity of the current selection battery (i.e. APM, SP100, SP200, SP301, SP302 and SP1600) in terms of academic performance of the students in the bridging programme.

The first specific empirical aim of this study was to describe the relationships between the applicable psychometric tests. The results obtained from investigating the correlations between independent variables showed that the APM statistically and practically significantly correlated (large effect) positively with both SP100 and SP400, as well as with SP301 (medium effect). The APM also shows a statistically significant relationship with the SP1600. The SP100 is significantly related to the SP400 (practically significant, medium effect) and statistically significantly related to the SP301. The SP200 is significantly and practically related only to the

SP301 (medium effect) and statistically significantly related to the SP302 and the SP1600. The SP301 is statistically significantly related to the SP1600. The SP302 is statistically significantly related to the SP400 (practically significant, medium effect) and the SP1600.

The second specific empirical aim of this study was to determine whether the instruments show predictive validity in this environment. The results of this study show that the APM is the test from the selection battery which correlates the highest with the Matric results. The APM, SP100, SP301 and SP400 are the only tests which show a statistically significant relationship with the Matric results when the whole battery was analysed. When these selected tests were analysed, not one showed statistically significant predictive power. The beta weights support the conclusion that the APM is the strongest predictor of the Matric results. In the light of this there is a level of predictive validity, although it is not as strong as was anticipated.

The third specific empirical aim of this study was to determine to what extent the applicable instruments predict/declare the variance in Matric performance. The results show that only the APM correlated positively with the Matric results, while all the selected SpEEx tests individually do not appear to be good predictors of the students' Matric results. The total selection battery has no statistically significant linear relationship with the dependent variable and it explains only 5% (adjusted R^2) of the variance in the dependent variable. Thus, only a small part of the variance of the dependent variable is explained by the independent variables. The total selection battery appears to be a poor predictor of the Matric results.

The results thus showed that although the current selection battery used for the bridging students does have some value, it only appears to be a poor predictor of the Matric results. All the SpEEx tests in the battery individually do not appear to be good predictors of the Matric results, while the respective beta weights of the individual instruments confirm that the APM is the strongest predictor.

The APM provides an indication of a person's intellectual efficiency (e.g. speed of accurate intellectual work), which is a critical skill required in the success of the Matriculation examination. This explains the positive correlation between this test and the Matric results (dependent variable) of the students.

A possible reason for the poor correlation of the SpEEx could be the restrictive range of items (the majority of the SpEEx tests only consist of 20 items) and thus does not discriminate sufficiently. Too few items lead to a small variance, which results in a positive skew curve. Tests with more items have a better distribution.

The value of the study was its confirmation that a positive relationship exists between the APM instrument in the selection battery and the university bridging students' Matric results.

5.2 LIMITATIONS

The fourth specific empirical aim was to discuss the limitations of this study. The identified limitations with respect to this study are listed below in terms of the literature review and empirical study:

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

- Limited research has been conducted on the selection of the students for the University Bridging Programme.

- The assumption was made that the final examination of the candidates (which was used as the dependent variable in this study) was reliable and valid.

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

- 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 concerning other culture and gender groups, e.g. women and white candidates.

- The majority of the sample group had other home languages, only five having English as a home language. All the tests were administered in English; this could have affected the scores in some way.

- The data was not analysed at item level because all the psychometric tests which were used are registered at the HPCSA as classified psychometric tests. They also all adhere to the requirements set by the Psychometric Board in terms of standardisation reliability and validity requirements.

5.3 RECOMMENDATIONS

The final specific aim of this study was to also formulate recommendations based on the current study in order to improve the current selection battery. Taking the above-mentioned limitations into account, the following recommendations for future studies are made:

- A comprehensive, up-to-date database pertaining to the university bridging students and their Matric results should be designed and maintained.

- Validation studies should be undertaken on a regular basis as the University Bridging Programme progresses and the population of these students increase over the years.

- The APM, which shows a positive correlation with the students' Matric results and thus seems to be the most valuable test in the battery, should be retained as part of a new university bridging student selection battery to predict success for previously disadvantaged students.

- The use of the SpEEx tests as part of the new university bridging student selection battery should be revisited.

- The current selection battery only measures specific abilities of the bridging student. The possibility of including instruments which measure other aspects, such as personality and interest, should be considered.

- It is suggested that clinical instruments (motivation testing) should also be used in the selection battery.

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

5.4 SUMMARY

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

Alderton, D.L., & Larson, G.E. (1990). Dimensionality of Raven's Advanced Progressive Matrices items. *Educational and Psychological Measurement*, 50, 887-900.

American Education Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological tests*. Washington, DC: American Psychological Association.

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.

Arthur, W., & Day, D. (1994). Development of a short form for the Raven Advanced Progressive Matrices test. *Educational and Psychological Measurement*, 54(2), 395-403.

Babbie, E. (1992). *The Practice of Social Research (6th ed.).* London: Thompson Publishing Company.

Bendix, S. (1996). *Industrial Psychology in the new South Africa.* Cape Town: Juta.

Blake, R.H. (1983). *Personnel Selection: Behaviour in organisation: South Africa perspective*. Johannesburg: McGraw-Hill.

Bors, D.A., & Stokes, T.L. (1998). Raven's Advanced Progressive Matrices: Norms for first year University Students. *Educational and Psychological Measurement*, 58(3), 382-398.

Brown, F.G. (1983). *Principles of Educational and Psychological testing* (3rd ed.). USA: Holt, Rinehart and Winston.

Cascio, W.F. (1991). *Applied Psychology in Personnel Management.* Englewood Cliffs: Prentice-Hall.

Dane, F.C. (1990). *Research Methods.* Pacific Grove, California: Brooks/Cole Publishing Company.

DACST (1996). *South Africa's Green Paper on Science and Technology. Preparing for the 21st Century.* The Department of Arts, Culture, Science and Technology, Pretoria, January 1996.

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).

Erasmus, P.R., & Minnaar, G.G. (1995). Realising human potential: the crux of the RDP. *Human resource management* 11(1).

Erasmus, P.E. (2001) *Situation Specific Job Profiling and Assessment in the workplace of the 21st Century.* Johannesburg Potential Index Associates cc.

Gerber, P.D., Nel, P.S. & Van Dyk, P.S. (1998). *Human resources management.* (4th ed.). Halfway House: International Thomson.

Ghiselli, E.E. (1964). *Theory of psychological measurement.* New York: McGraw-Hill.

Gregory, R.J. (1996). *Psychological testing: history, principles, and applications* (2nd ed.). Boston: Allyn & Bacon.

Helmstadtler, G.C. (1964). *Principles of Psychological Measurement.* New York: Meredith Publishing Company.

Herholdt, W. (1977). *'n Keuringsprogram vir Argitekstudente.* Unpublished Doctoral Thesis. University of Pretoria.

Howell, D.C. (1995). *Fundamental statistics for the behavioral sciences.* (3rd ed.). Belmont: Duxbury Press.

Huysamen, G.K. (1983). *Psychological Measurement: An introduction with South African examples.* Pretoria: Academia.

Huysamen, G.K. (1994). *Methodology for the Social and Behavioural Sciences.* Cape Town: Southern Book Publishers.

Jacobs, C.D., Haasbroek, J.B. & Theron, S.W. (1992). *Effektiewe Navorsing.* University of Pretoria.

Jordaan, W.J., & Jordaan, J.J. (1986). *Man in context* (3rd ed.). Johannesburg: McGraw-Hill.

Kerlinger, F.N., & Lee, H.B. (2000). *Foundations of Behavioural Research.* (Fourth Edition) Fort Worth: Harcourt College Publishers.

Klinvex, B. (1999). *Organizational psychology.* New York: Macmillan.

Labour Relations Act, No 66 (1995). *Government Gazette, 366* (16861).

McLaurin, W., Jenkins, J., Farrar, W., & Rumore, M. (1973). Correlations of IQ on verbal and non-verbal test of intelligence. *Psychological Reports*, 33, 821-822.

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.

Mouton, J., & Marais, H.C. (1994). *Basic Methodology of the Social Sciences.* Pretoria: Human Sciences Research Council.

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.

Raven, J.C. & Court, J.H. (1985). *Manual for Ravens Progressive Matrices.* London: H.K. Lewis.

Raven, J., Raven, J.C., & Court, J.H. (1998). *General Overview. Raven Manual: Section 1.* Oxford, UK: Oxford Psychologists Press.

Raven, J., Raven, J.C., & Court, J.H. (1998a). *Advanced Progressive Matrices. Raven Manual: Section 4.* Oxford, UK: Oxford Psychologists Press.

Reber, A.S. (1988). *The penguin dictionary of psychology.* London: Penguin.

Rosnow, R.L., & Rosenthal, R. (1996). *Beginning behavioral research: A conceptual primer (2nd ed.).* Englewood Cliffs: Prentice-Hall.

Rudner, L.M. (1994). *Questions to ask when evaluating tests.* Washington, DC: ERIC Clearinghouse on Assessment and Evaluation.

Rushton, J.P. (2004). Construct Validity of Raven's Advanced Progressive Matrices for African and Non-African Engineering Students in South Africa, *International Journal of Selection and Assessment,* Vol 12(3), 220-229.

Selltiz, C., Jahoda, M., Deutsch, M., & Cook, S.W. (1976). *Research methods in social relations.* Revised edition. New York: Holt, Rinehart & Winston.

Schaap, P. (1997). *The predictive and construct validity of the PIB in a financial institution.*

Pretoria: University of Pretoria.

Schaap, P. (2001). *The Psychometric Properties of the SpEEx.* Pretoria: University of Pretoria.

Schaap, P. (2001a). *The Psychometric Properties of the SpEEx (An Updated Version).* Pretoria : University of Pretoria.

SIOP. (2003). *Principles for the Validation and Use of Personnel Selection Procedures.* Fourth Edition. www.siop.org.

Smit, G.J. (1996). *Psychometrics: aspects of measurement. Pretoria: Kagiso.*

Society for Industrial Psychology. (1998). *Code of practice for psychological assessment in the workplace.* Pretoria: SIP.

Society for Industrial Psychology. (1998). *Guidelines for the Validation and use of Assessment Procedures for the Workplace.* Pretoria: Society for Industrial Psychology.

SPSS. (2005). *SPSS 14.0.* Chicago, IL: SPSS Inc.

Tabachnick, B.G. & Fidell, L.S. (1996). *Using Multivariate Statistics* (3rd ed.). Northridge: Harper Collins College.

The Constitution of the Republic of South Africa, 1996. (Act No 108 of 1996).

Walsh, W.B. & Betz, N.E. (1995). *Tests and assessment* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.