

CHAPTER 1

BACKGROUND FOR AND MOTIVATION OF THE RESEARCH

1.1 INTRODUCTION

The influx of students from previously disadvantaged groups to tertiary institutions places a renewed emphasis on selection of students. Some students are not ready for tertiary education, because of their poor educational background (Huysamen, 2001; Miller, 1992) and therefore the potential of students to perform at a required level to obtain a qualification needs to be identified. Malekele (1994, p. 1) points out that “tertiary education has developed along ‘separate’ and ‘unequal’ lines of primary and secondary education”.

Continued research to investigate the validity of selection batteries for predicting academic success at tertiary institutions, has been done over decades and still needs to continue (Huysamen, 1999, 2001; Malekele, 1994; Van Eeden, De Beer & Coetzee, 2001; Van Rooyen, 2001). In the United States of America, predictive validity studies are done on a regular basis by the Educational Testing Services (Huysamen, 2000). Malekele (1994) notes that countries like the USA and Australia successfully use psychometric assessment for student selection, but expresses his concern regarding the fairness of psychometric tests used in South Africa. Lewis, Morgan and Ramist as cited by Huysamen (2001), also recommend regular validity studies at tertiary institutions in South Africa.

To understand the need for this study one can question the appropriateness of including psychometric tests as part of a selection battery at tertiary institutions. Psychometric tests are mainly included as part of a selection battery due to the following three reasons:

- The lack of predictive validity of matric results (The term matric will be used throughout this document as synonym for what is currently known as the Grade 12 or Senior Certificate, as it was still known as matric when the data of this project were gathered).
- The necessity of fair selection in a multicultural society.
- Scarcity of funds and resources at institutions of higher education.

The first reason for including psychometric testing in a selection battery is that there is little evidence of a correlation between matric results and academic performance (Huysamen, 2000; Miller, 1992). The correlation problem is further complicated by educational inequalities (Huysamen, 2001; Malekele, 1994; Miller, 1992). Matric results alone can therefore not be used as a valid and reliable predictor of academic performance. Miller (1992) pointed out that it would be unfair to have students coming from different school backgrounds competing for admission on equal footing. Miller (1992), as well as Van Eeden et al. (2001), indicates that it is necessary to use a combination of matric results and psychometric tests in a selection battery, to increase the predictive validity of the selection process.

Secondly, the moral concern for fair selection of suitable and deserving candidates in South Africa necessitates the inclusion of psychometric testing in a selection battery. In this regard the kind of psychometric test to be used becomes an issue. Boeyens (1989a) and Jonker (1994) suggested that, to ensure fair selection, provision should be made for students who had inadequate school education, but have the *potential* to obtain a tertiary qualification. The psychometric tests that are used for selection purposes should therefore assess learning potential rather than measuring crystallised cognitive competencies that often reflect formal education and school background.

All psychometric tests are not necessarily fair to use in a selection battery for people from different cultures, different socio-economic backgrounds or different educational backgrounds. Students from a low socio-economic group and/or educational

disadvantaged background usually score lower marks on achievement and aptitude tests (Claassen, 1990a; Jensen, 1980; Owen, 1998). In competing for admission to tertiary institutions, these applicants are at a disadvantage if they compete with candidates who have a better educational background (Huysamen, 1996a; Rutherford & Watson, 1990). The Health Professions Council of South Africa appealed for the development of culturally appropriate measures (De Beer, 2000c).

The psychometric tests used should ensure that students coming from previously disadvantaged schools will have a fair opportunity to compete with students coming from historically white schools for the limited positions available at institutions of higher education (Maxwell, 1998; Nunns & Ortlepp, 1994; Zaaiman, 1998). A high emphasis is therefore placed on the re-evaluation of the appropriateness of traditional psychometric testing, which supports the need for this study.

The high costs of tertiary education and the scarcity of state funds necessitate that only students with a reasonable chance to successfully complete their studies should be identified through a selection process and admitted to a tertiary institution (Huysamen & Raubenheimer, 1999; Jonker, 1994). Kotze (1994), Samkin (1996), as well as Stoker, Engelbrecht, Crowther, Du Toit and Herbst (1985), are some of the authors who reported on the unacceptably high failure rates of students and investigated alternative selection criteria that are more accurate predictors of future academic performance. Uys (1994) referred to students who fail as wastage of funds and stated that the economic loss for the state is significant.

Local research to determine whether psychometric tests can predict academic performance for all students at tertiary institutions, regardless of their socio-economic and school background, is therefore imperative (De Beer, 2000c; Malekele, 1994; Miller, 1992).

The above explanations provide some reasons for the need to investigate the use of psychometric tests, together with matric results, as part of a selection battery for the admission of students to a tertiary institution.

1.2 PROBLEM STATEMENT

Little research was found, with regard to investigating the predictive validity of the particular psychometric tests used for the selection of students in commercial courses at technikons, specifically at a distant campus (satellite campus). Only the studies of Ayaya (1996) and Huysamen (2000), regarding the academic performance of students over the full duration of their studies, were found as examples of research in this regard in South Africa.

According to Ayaya (1996), a question can be raised about the predictive validity of the admission measures for the full duration of the qualification. Wilson (1983) points out that high predictive validity for first-year academic performance alone does not represent a sufficient sample of a student's academic performance for the full duration of a course. There is evidence of improved academic performance from the second academic year onwards, which is referred to as the late-blooming hypothesis (Huysamen, 2000). Based on the above, this research project was conducted over a three-year period, in order to obtain academic information of students in the minimum time to complete a national diploma.

The current selection battery that is used at the Technikon Pretoria Nelspruit Campus consists of a combination of certain sub-tests of the Potential Index Battery (PIB) and matric results. Poor pass-rates in the major subjects indicate that the predictive validity of the current selection battery needs to be investigated. This implies an investigation of the correlation between the academic performance, matric results and scores obtained in the selection tests that are currently used to admit students.

The sample of students used in the study was also exposed to a dynamic psychometric instrument, namely the Learning Potential Computerised Adaptive Test (LPCAT), because of the need to measure learning potential rather than crystallised intelligence. The LPCAT is a newly developed test that aims to assess the learning potential of individuals (De Beer, 2000c).

1.3 AIMS

1.3.1 General aim

The general aim of this research is to ascertain the predictive validity of a selection battery for technikon students using their academic performance over the duration of three years as criterion measure.

1.3.2 Specific aims

The following specific aims will be addressed in this study:

- to provide a literature review of theoretical concepts like intelligence, learning potential, student selection and changes in assessment of cognitive abilities.
- an empirical investigation of the academic results of students over a three year period.
- determining the correlation between matric results and academic performance.
- investigating whether any specific matric subject shows a statistically significant correlation with academic performance.
- to determine the predictive validity of two psychometric tests, in terms of the correlation of each psychometric test with academic performance, after each academic year and over the full duration of the course (three years).
- a comparison of the predictive validity of the two psychometric tests.

- to determine the correlation between academic performance in the national diploma and the bridging course status of students.
- to determine whether actual scores obtained in the first year of study significantly correlate with academic performance in the second and third academic years.

1.4 THE PARADIGM PERSPECTIVE

1.4.1 The disciplinary relationship

This research project falls within the psychometric discipline which is a sub discipline of industrial and organisational psychology.

Psychometric testing is concerned with the measurement of different aspects of human behaviour and potential, through the development of psychological tests (Anastasi & Urbina, 1997). It is also concerned with the development of new techniques for statistical analysis (Weiten, 1994).

This study is relevant for industrial psychology, because proper selection of students in terms of their interests, abilities and potential, is the first step towards successful integration of individuals with their future careers. The use of dynamic psychometric measures can also assist in identifying individuals with potential, which can be helpful in addressing equity and imbalances of the past.

1.4.2 Applicable psychological paradigms

This study draws from the functionalistic, behaviouristic and cognitive schools of thought in psychology.

The early functionalists, James and Dewey, focused on the way an organism adapts to its environment and investigated how the mind functions (Papalia & Olds, 1988). Scientific methods of data collection and the use of questionnaires and mental tests, gained from this school of thought, are still used in psychology today (Baron, 1994; Papalia & Olds, 1988; Weiten, 1994).

According to Baron (1994), the behaviourists, under leadership of Watson, focused more on observable behaviour and events. The nature of their research moved towards experiments with animals and learning. They believed that behaviour is measurable and studied stimulus-response relationships. The behaviourists investigated the role of the environment in shaping human nature and indicated that it is not only the impact of heredity factors that influences cognitive abilities of people. A major objection against behaviourism is its denial of the cognitive processes of humans. Since the early years in the study of psychology, theorists and researchers have argued about the influence of heredity and the environment on cognitive development (Baron, 1994; Papalia & Olds, 1988; Weiten, 1994). There are indications that both heredity and the environment influence intelligence (Anastasi, 1988; Claassen, 1990a; Baron, 1994; Hamers & Resing, 1993). There are, however, still different views on which one plays a bigger or more important role (Anastasi & Urbina, 1997).

Lastly, the cognitive school of thought, which investigates processes of the mind, specifically the way in which the mind perceives, organises, remembers and uses information, applies to this study (Papalia & Olds, 1988).

1.4.3 Applicable behaviour models

The literature review that will be reported on in chapter two will focus on the evolution of measures of cognitive abilities, focusing on the traditional models since Binet and Simon and other theorists (Baron, 1994; Papalia & Olds, 1988; Weiten, 1994) as well as more

recent models measuring learning potential, like the model of Vygotsky (Vygotsky, 1978) and other supporters of dynamic assessment.

In South Africa it is essential to also consider potential future performance, in other words, to assess learning potential, rather than assess only current abilities. The Vygotsky model is specifically suitable to the current situation in South Africa, where there are educational inequalities and differences in socio-economic backgrounds. This model allows for improvement of cognitive performance if relevant training is provided (De Beer, 2000c; Vygotsky, 1978).

1.4.4 Applicable concepts and constructs

The concepts and constructs that will be investigated in this study are mainly cognitive ability and learning potential. The terms cognitive ability and intelligence are used interchangeably in the literature (Anastasi & Urbina, 1997) and they will also be used interchangeably in this document.

Intelligence is defined by Papalia and Olds (1988) as a constant interaction between inherited ability and environmental experience, which results in a person's ability to acquire, remember, and use knowledge; to understand both concrete and abstract concepts; to understand relationships among objects, events, and ideas; and apply and use all these abilities in a purposeful way to solve problems in everyday life. Various other definitions of intelligence are available, but this definition will be used as it illustrates that both heredity and environment impact on intelligence.

Learning potential is defined by De Beer (2000c, p. 15) as “an overall capacity and includes both present and improved future performance”. De Beer (2000c) further indicates that assessment of learning potential emphasises development and allows for improvement in cognitive performance, if relevant training is provided, in other words focusing on changing the level of functioning of the individual. Learning potential

measurement refers to unrevealed cognitive capacities according to Feuerstein, Feuerstein and Gross (1997).

The measurement of learning potential is therefore different from traditional measures of cognitive ability. Traditional tests measuring cognitive ability only provide an estimate of current performance, while tests measuring learning potential give an indication of potential future performance. More definitions of these concepts will be provided in chapter two.

The reference to matric in this project refers to what is currently known as Grade 12 which is the final year of secondary education. The term matric will be used throughout this document as students still matriculated in the year that the data were collected.

1.4.5 Methodological convictions

Psychometric tests are used in this project to measure cognitive abilities of people. A wide variety of psychometric tests are available. The major categories are tests of cognitive abilities, tests of specific abilities and personality tests (Anastasi, 1988). According to Anastasi (1988, p. 23) “A psychological test is essentially an objective and standardised measure of a sample of behavior”. Cognitive abilities are regarded as a measurable construct (Papalia & Olds, 1988). Anastasi (1988, p. 24) states that “predictive validity of a psychological test depends on the degree to which it serves as an indicator of a relative broad and significant area of the behavior”.

Correlation analysis and regression analysis are the two statistical methods that were used to quantify and describe the relationship between the dependent and independent variables and to determine the predictive power of the independent variables for the dependent variables.

1.5 RESEARCH DESIGN

Mouton (1998) compares research with the planning of a journey. According to him research design results in a step-by-step process of how the research project will be conducted to ensure maximum validity in the study.

This research project is a longitudinal quantitative study that investigates the relationships between dependent and independent variables. Variables are related to each other; when one variable changes, the other variable will also change. Independent variables are also known as predictor measures of a research project. They reflect the presumed cause of the dependent variable, while a dependent variable or criterion measure refers to effect or result of the independent variable (Mouton, 1998; Tredoux & Durrheim, 2002; Wegner, 1993). The independent research variables in this study are the matric results and the scores that students obtained in two psychometric tests (PIB and LPCAT), as well as data with regard to students' bridging course status and gender. The dependent variable in this research project is the tertiary academic performance of students over the three-year period of their studies.

Registered students at the Nelspruit satellite campus of Technikon Pretoria between 2000 and 2002 were used for this study. Only students who registered for the National Diploma: Human Resources Management for the first time in 2000 were included in the sample. Academic records of students were collected over a three-year period. Biographical data with regards to the age, gender and race were obtained from the students' files. Although the data of students are obtained on individual level, the correlation will be determined for the group results.

Standardised psychometric tests were used in the study to ensure valid and reliable results. Examination results are regarded as objective measures of the students' knowledge and understanding of each subject, and are therefore regarded as valid and reliable data. The academic performance of the students was recorded over a three-year period, which is the full duration of the National Diploma course.

The data gathered over the three-year period were statistically analysed to obtain the correlation and regression values between the variables. This information will be used to report on the predictive validity of the selection battery used for the selection of students. The strength and statistical significance of the correlations and regression analysis will indicate the extent to which matric results, the two psychometric tests and/or other predictors are valid predictors of academic performance.

1.6 CHAPTER ALLOCATION

The following lay-out of chapters is used:

This chapter has been an introductory chapter and included the background of the research topic and the purpose and procedures of the study.

A report on the literature review providing information on existing theory and research results, related to the field researched in this project, will follow in chapter two. The measurement of cognitive abilities will be explained. In chapter two specific focus is placed on the new dynamic tests, as well as the importance of culture fairness in psychometric testing.

In chapter three the theoretical overview is continued and student selection at tertiary institutions in South Africa is discussed.

In chapter four the research process and design of the study will be explained. The two psychometric tests used as measuring instruments in the study will also be discussed.

The results and statistical analysis of the data are included and discussed in chapter five.

In chapter six the results will be discussed in terms of the research problem and the findings will be integrated with the literature review. In this last chapter the limitations will be highlighted and recommendations will be made.

CHAPTER 2

COGNITIVE ABILITY AND ITS MEASUREMENT

2.1 INTRODUCTION

The measurement of cognitive abilities is usually included in a selection battery and is used as a predictor of academic performance. There has been a lot of criticism regarding the suitability of traditional static psychometric tests, as a measure of cognitive ability, in the South African context. As a result of the criticism, preference for the use of dynamic assessment instruments has developed (Boeyens, 1989; De Beer, 2000c; Shochet, 1986; Van Aswegen, 1997).

In this chapter the various views and definitions of cognitive abilities as well as the nature nurture debate will be reviewed to reflect on the effect of heredity and the role that the environment plays in the measured level of cognitive ability of a person.

The use of psychometric tests to measure cognitive abilities, will be discussed, with reference to the requirements to ensure valid and reliable psychometric test results. Thereafter, the challenges of psychometric testing will be reviewed.

The different approaches in measuring cognitive abilities will be discussed. Specific attention will be given to traditional measures of intelligence and the more modern dynamic methods, with an indication of where the two psychometric tests used in the study fit in.

2.2 DESCRIBING THE CONCEPTS

Prior to reporting on the evolution and measurement of cognitive abilities, a few terms need to be clarified.

2.2.1 Cognitive ability

Cognitive and mental ability are used as synonymous terms for intelligence in the literature (Tuckman, 1975) and will also be used interchangeably in this project. There are various views and theories on intelligence, which have resulted in different definitions of intelligence. The definition of intelligence depends on the theoretical view that is held (Anastasi, 1988). The different views are discussed in section 2.3.

Owen (1998, p.17) indicated that “Most people can provide a general description of intelligent behaviour, but experience problems as soon as a more particular definition is required”.

According to Eysenck (1988), the term intelligence was traditionally used in three different senses, namely biological intelligence, psychometric intelligence and social or practical intelligence. Biological intelligence is represented by Galton’s view on intelligence, which focused on the structure of the human brain, its physiology, biochemistry and genetic aspects. Psychometric intelligence, or the intelligence quotient (IQ), refers to intelligence as measured by standardised psychometric tests and is generally seen to be determined by biological and cultural factors, family upbringing, socio-economic status and education, amongst others (Anastasi & Urbina, 1997). Social or practical intelligence refers to the social competencies of humans and the application of IQ to solve the problems an individual encounters (Eysenck, 1988; Sternberg, 1985).

2.2.2 Learning potential

Hamers and Sijtsma (1993) stated that learning ability or learning potential is equivalent to the simplistic definition of intelligence, viewing it as the ability to learn, while traditional IQ tests measure current cognitive abilities.

Babad and Budoff (1974) describe learning potential as the trainability, that is, the ability of an individual to improve performance in reasoning problems, after a learning experience.

De Beer (2000c, p. 105), who developed a test for the measurement of learning potential in South Africa, defined learning potential as “the ability to benefit from (new) learning experiences by using appropriate existing and realisable skills” and views it as a measure of general ability.

2.2.3 Dynamic assessment

Dynamic assessment refers to the dynamic nature of the testing process where the administering of the test involves a test, train and re-test process, while traditional cognitive tests is static in nature with no intervention during the testing process (Boeyens, 1989b; Cronbach, 1990; De Beer, 2000c; Feuerstein, Rand, Jensen, Kaniel and Tzuriel, 1987; Guthke, 1992; Hoy & Gregg, 1993; Murphy, 2002; Owen, 1998).

De Beer (2000c) and Guthke (1992) indicated that there are various terms that are used to describe modern cognitive tests. Some of these equivalent terms are “cognitive enrichment”, “learning tests”, “dynamic assessment or testing”, “testing the limits” and “learning potential measurement”.

There are mainly two approaches to dynamic assessment. They are the enrichment or remedial approach and the measurement or psychometric approach. The enrichment approach focuses on structural changes resulting from remedial mediation, while the measurement approach focuses on measuring the changes in performance (Grigorenko & Sternberg, 1998). The dynamic assessment instrument used in this project resorts under the measurement approach of dynamic assessment. More attention will be given to these different approaches in section 2.8 and section 2.9.

2.3 THE EVOLUTION OF THEORIES ON COGNITIVE ABILITIES

Anastasi (1988), Brown and French (1979), Gregory (1986), Jensen (1980) and Sternberg (1985) note that psychologists cannot agree on a definition of intelligence and that this complicates measurement. Eysenck (1988) indicated that although there was controversy regarding the nature and existence of intelligence, its measurement was still regarded as one of the greatest achievements in psychology in the 'between wars years'. Doubts have been expressed about the existence of intelligence, its measurement and the meaningfulness of traditional measures of the intelligence quotient (IQ). According to Sternberg (1985), intelligence is still regarded as an elusive concept. The different theories that exist view intelligence as either a unified or multifaceted construct, in other words, either as a single characteristic or several distinct parts (Anastasi, 1988; Anastasi & Urbina, 1997).

Plato was the first to associate the brain with mental processes (Verster, 1983). During the middle ages, Plato suggested a hierarchy of the mind and referred to three components of the human mind, namely the intellect, emotion and the will of an individual (Claassen, 1990a; Gardner, 1993). The term intelligence comes from the work of Aristoteles, one of the students of Plato. Cicero translated the work of Aristoteles and was the first to create the Latin word "*intelligentia*", which means intelligence (Claassen, 1990a; Gardner, 1993; Verster, 1983).

The first attempts to measure intelligence were made at the time when Wundt managed a psychological laboratory at Leipzig University, Germany, in 1879, where experimental psychology developed. He studied mental events through introspection (Gardner, 1993; Hilgard, 1989; Spearman, 1904).

Gardner (1993) and Hilgard (1989) indicate that Gall, a German anatomist, directed attention to the physical differences of the mind, as correlates of differences in psychological faculties of the mind. We know today that the size of the brain has no correlation with the intellect of the individual (Gardner, 1993).

Galton, who established a laboratory in London in 1882, measured individual differences in sensory functioning through reaction time studies. Galton was regarded as the father of mental testing (Anastasi, 1988). He regarded reaction time as an indicator of intellectual capabilities and believed that intelligence could be measured accurately (Boeyens, 1989a; Brody, 1992; Cronbach, 1990; Hilgard, 1989; Spearman, 1904; Taylor, 1994). Galton viewed individual differences as objectively measurable by standardised procedures (Gregory, 1986). His work resulted in the recognition that human traits are normally distributed (Owen, 1998; Verster, 1983). Tests based on Galton's view of intelligence, measuring broad-based psychological constructs such as abilities, are still in use today (Taylor, 1994).

Alfred Binet, a French researcher, began his investigations of individual differences by observations of his own daughters (Brody, 1992; Tuckman, 1975). Binet worked with various researchers and spent a lot of time researching ways of measuring intelligence (Biesheuvel, 1943; Brody, 1992; Eysenck, 1971; Gardner, 1993). Theodore Simon later joined him and they published the first intelligence test in 1905 on instruction of the French school authorities (Claassen, 1990a). Translations and adaptations of their test were used throughout the world (Anastasi, 1988). This was the first test to use the term intelligence quotient (IQ). Terman revised the Binet scales in 1916 and the Stanford–Binet scale for American children was developed. This test was revised as recently as 1986 and is still in use. Hilliard (1990) pointed out that, according to Gould (1981), Terman not only translated Binet's test, but made major changes to it.

Binet was the first to establish the concept “mental age”. Mental age, according to Biesheuvel (1943, p. 13), indicates “the degree of mental development, which one can expect from a normal or average child at a particular age level”. Chronological age represents the real or actual age of a person. The IQ score is calculated by dividing the mental age by the chronological (or actual) age and then multiplying it with 100 (Brody, 1992; Hilgard, 1989; Mattarazzo, 1990).

Binet's work is significant to this project, as he was one of the first theorists to promote the idea of modifiability of intelligence and the measurement of the ability to learn. He viewed intelligence as changeable (Grigorenko & Sternberg, 1998; Hilliard, 1990).

Spearman was a strong supporter of the view that intelligence comprises a general unified capacity (Baron, 1994; Gardner, 1993; Spearman, 1904; Sternberg, 1985; Tuckman, 1975) and postulated a two-factor theory on intelligence. The two factors were the general intelligence, called the "g" factor and a second factor, consisting of several other independent specific factors, that he termed "s" factors (Anastasi & Urbina, 1997; Cronbach, 1990; Gardner, 1993; Horn, 1989; Tuckman, 1975; Verster, 1983). Spearman's theory was based on the principles of Galton (Gardner, 1993; Horn, 1989). He argued that "g" could represent individual differences in terms of different levels of mental energy. He also developed the correlation coefficient, a statistical technique that reflects the relation between two variables (Gardner, 1993).

Another theorist of intelligence, Cattell, believed that intelligence consists of two components, which he called fluid abilities and crystallised abilities. They represented Spearman's "g" (Brody, 1992; Cronbach, 1990; Gardner, 1993; Hilgard, 1989). Cattell was a student of Wundt, but was also influenced by Galton. Fluid abilities (g_f) are regarded as a basic inherited capacity related to novelty and flexible thinking and not influenced by the environment, while crystallised abilities (g_c) represent the storage and use of declarative knowledge, such as vocabulary. Crystallised abilities therefore are influenced by schooling and represent specialised skills and knowledge, promoted by the culture of the individual (Brody, 1992; Cronbach, 1990; Sternberg, 1985, 1997; Taylor, 1994).

Cattell was interested in the measurement of individual differences (Anastasi, 1988) and in 1940 was one of the first theorists who aimed to develop, what he initially called, culture free tests. The term "culture free" caused a lot of misunderstanding and after Cattell was accused by other psychologists of setting an illusionary goal, he later changed the term to "culture fair" intelligence tests (Jensen, 1980). Cattell's culture fair tests

aimed to measure fluid, rather than crystallised abilities (Taylor, 1994). Cattell further developed a causal analysis of the relationship between g_f and g_c . According to Cattell, factor g_f is more likely to influence g_c than the other way around. This results in a fundamental distinction between ability and achievement. In his view intellectual ability therefore indicates that one has the ability or the potential, which might not have been actualised. Achievement, on the other hand, indicates the attainment of something, for which ability is a pre-requisite. This indicates that it is not possible to acquire knowledge if there is no opportunity to learn. This relationship between abilities and achievement is complex and changes over time in Cattell's view. A person's achievements in life are therefore not only a function of the abilities of a person, but environmental circumstances can also impact on the achievement of a person (Brody, 1992).

Sternberg's triarchic theory is one of a few that represent the multifaceted view of intelligence with a broader conceptualisation of intelligence (Sternberg, 1985). In Sternberg's view intelligence is "a mental activity directed toward purposive adaptation to, and selection and shaping of, real world environments relevant to one's life" (Sternberg, 1985, p. 45). This is similar to the definition of social or practical intelligence (Eysenck, 1988).

Sternberg (1985, p. 46) goes further in saying:

because what is required for adaptation, selection, and shaping may differ across persons and groups, as well as across environments, *intelligence is not quite the same thing from one person (or group) to another, nor is it exactly the same thing across environments. Nor is intelligence likely to be exactly the same thing at different points in the life span.* [italics added]

It is clear from this statement that Sternberg views intelligence as a dynamic entity of an individual. It is also obvious that Sternberg believes that cognitive abilities are not

necessarily the same for different people or cultures. Sternberg's statement also indicates that the environment can have an effect on the intellectual capacities of people.

With regard to the measurement of cognitive abilities, Sternberg (1985) feels that the suitability of an instrument varies within and between socio-cultural groups and that it is best to use a battery of tests. He expressed himself against traditional psychometric tests, as he viewed them as only providing information on the performance of an individual in an intelligence test and not providing information on the ability of the individual to cope in the everyday world (Sternberg, 1985). Sternberg (1985, 1997) argues that theories of intelligence have been constrained and dominated by available measures of intelligence. He criticised theorists for first developing measures and then developing theories, while it should have been the other way around (Sternberg, 1985, 1997).

During the 1980's, Gardner, like Sternberg, viewed intelligence more as a multiple construct than a unitary construct (Gardner, 1993; Hamers & Resing, 1993; Sternberg, 1997). Gardner was concerned about the biased view of intelligence, focusing mostly on "crystallised" rather than "fluid" intelligence. According to Gardner, one must study the interaction between the individual and the everyday environment to be able to understand intelligence (Hamers & Resing, 1993). He argued that intelligence tests rarely assess skill in assimilating new information or solving new problems. He further argued that intelligence tests say little about an individual's potential for growth (Gardner, 1993).

Jean Piaget emphasised the adaptive nature of intelligence. The Piagetian approach describes the qualitative changes in thinking that are possible (Papalia & Olds, 1988). Piaget viewed intelligence as a biological process, with a focus on the development of a set of cognitive structures during the four stages of development from birth (Brown & Campione, 1986; Owen, 1998).

These four stages are:

- birth – 24 months - This is called the sensori-motor stage, when the child becomes aware of objects and how they differ from the self.
- ± 24 months – 7 years - This is called the pre-operational stage, during which a child understands the cause and effect of relationships.
- ± 7 – 11 years - This is called the concrete operations stage when the child develops the ability to think independently.
- from 11 years onwards - This is called the formal operations stage, when the young child can reason logically.

Intelligence can be modified during these stages, or even limited by the environmental setting in which it occurs, according to Piaget (Claassen, 1990a; Retief, 1988).

Another paradigm, with regard to the modifiability of intelligence, is presented by Eysenck's model of intelligence, in which he questions the modifiability of intelligence (Eysenck, 1988). This theory regards the modifiability of intelligence and the speed of cognitive processing, as the fundamental variables underlying differences in general intelligence (Cronbach, 1990a; Eysenck, 1988).

As views on intelligence changed and developed, other types of intelligence were also identified, for example practical intelligence, emotional intelligence, social intelligence and creative intelligence (Gardner, 1993; Sternberg, 1985). These developments resulted in new kinds of cognitive measures, of which one is dynamic assessment (Sternberg, 1997). The dynamic assessment approach developed from the theoretical base of the

Russian psychologist, Vygotsky (Boeyens, 1989a; Campione & Brown, 1987; Vygotsky, 1978). The approach is based on the philosophy that guided learning provides a more accurate measure of potential than static measurement through conventional tests. This approach makes a distinction between the actual level of development, measured by tests using the standard psychometric approach, and the potential development, as determined by dynamic assessment, that measures performance following a training intervention (Feuerstein, 1973; Hoy & Gregg, 1993; Owen, 1998; Sternberg, 1997; Taylor, 1994; Vygotsky, 1978).

The dynamic assessment view of cognitive assessment is the principle on which one of the measuring instruments, used in this research project, is based. This concept will be discussed in more detail in section 2.8 and 2.9.

2.4 THE NATURE NURTURE DEBATE

A discussion of the views and theories of cognitive abilities will not be complete without reference to the nature nurture debate. Although differences in intellectual capacities between individuals are noticeable and do not have to be proved by tests, the reasons underlining these differences are less obvious. The nature nurture debate deals with the question whether the differences can be attributed to heredity or to environmental factors or a combination of both (Baron, 1994; Hamers & Resing, 1993).

To prove whether heredity plays a role, studies on families were conducted. Twin studies, where identical twins were raised together, provided the highest correlation between the IQ's of the two siblings. In cases where identical twins were raised apart, again, a high correlation between the IQ of the two siblings was found (Anastasi, 1988; Baron, 1994; Cronbach, 1990; Eysenck, 1971). This provides evidence for the role of genetics in the level of intellect.

Shochet (1986, p. 55) summarised the principles on which the genetic position is based as:

- The general intelligence factor, *g*, is the common factor of intelligence tests.
- Intelligence is predominantly biologically determined and 80 percent of intelligence is inherited.
- Differences in IQ scores are not a result of test bias, but due to different cultures' levels of intelligence.

Research was also conducted to determine the influence of environmental factors on intelligence. Some of the environmental factors that are considered to play a role are, amongst others, standard of living, socio-economic status (SES) of parents, quality of diet, and educational opportunities. Research on deprivation supports the idea that the environment plays a role (Baron, 1994). Studies on foster children, taken from orphanages, confirmed that these foster children show relatively higher IQ's than would be expected from their biological parental backgrounds. This indicates the advantages gained by living in better conditions, as foster homes are carefully selected by adoption agencies (Hilgard, 1989). The effect of the environment on intelligence is indicated by these studies.

One of the environmental factors that has a significant influence on intelligence, is the socio-economic status (SES) of the parents (Claassen, 1990a). Binet and Simon (1916), as cited by Claassen (1990a), already indicated that their test of 1905 was developed for the working class in France, and they found that children coming from a higher social class scored significantly higher in the test. Claassen (1990a) indicated that differences of 15 to 25 points on the scale could be expected from children with parents practicing a professional career, compared to children with parents employed in low literacy level jobs. This indicates the effect that socio-economic status can have on measured intelligence. Jensen (1969) also conducted research on children from groups with a low socio-economic status. He concluded that the nature nurture problem relates to socio-

economic status and emphasised the need for tests to assess a broader spectrum of mental abilities. This indicates the need for measurement of learning potential, rather than the traditional static measures of intelligence. Research done by Babad and Budoff (1974) also supported the findings of Jensen. They reported that learning potential scores of individuals with a low IQ, coming from groups with a low socio-economic status, showed ability to reason. This ability was not indicated by the results obtained from the same sample in standard tests of intelligence.

Based on the view that the environment can influence cognitive abilities, “the environmentalists ... have developed specific intervention and enrichment programmes in an attempt to raise the intelligence of individuals or large groups” (Hamers & Resing, 1993, p. 26). This is also the background against which dynamic assessment has developed.

There is strong evidence for both environmental and genetic views and most psychologists accept that both factors play a role in intelligence (Hamers & Resing, 1993).

2.5 PSYCHOMETRIC TESTING

Psychometric testing is used to measure unobservable characteristics, of which cognitive ability is one. According to Anastasi (1988, p. 336), psychological tests are like tools and the “effectiveness of the tool depends on the knowledge, skill and integrity of the user”. Owen (1998) indicates that the existence of psychological testing is a result of the differences between individuals. If people were the same, there would be no need for testing.

Many definitions of psychometric tests or psychological tests exist. Some of the definitions are indicated below:

- According to both Anastasi (1988) and Anastasi and Urbina (1997), a psychological test is essentially an objective and standardised measure of a sample of behaviour.
- Cronbach's (1990, p. 32) definition of a test is a "systematic procedure for observing a person's behaviour and describing it with the aid of a numerical scale or category system".
- Owen (1998) indicated that psychometric tests provide information about the typical behaviour of an individual.
- Seligman (1994) stated that psychological measurement is a source of diagnostic information about clients.
- De Beer (2000c, p. 28) defines psychological testing as "the use of psychometric instruments to obtain information about individuals or groups so that understanding of people concerned and decision making that affects them can be improved".

Testing involves measuring of behaviour displayed during the test. The most important objective is to distinguish between inter and intra-individual differences (Anastasi, 1988; Boeyens, 1989a; Owen, 1998). Anastasi (1988) pointed out that tests are designed to indicate what an individual can do at a certain time, but not to provide reasons for what the individual can or cannot do.

Erasmus (1997), one of the developers of the PIB, a psychometric test used in this project, prefers to talk about an assessment, rather than a test. He argues that "test" means that a person can pass or fail, while assessment can be regarded as a "mental fingerprint" of a person's cognitive ability. Cronbach (1990) also regards assessment as a wider concept than a test, as it includes integration and evaluation of information.

Owen (1998) highlighted the difference between measurement and evaluation. According to him, measurement refers to the process of determining and allocating a score to characteristic features through observation and testing. The administering of a test can be regarded as the measurement process. Evaluation, on the other hand, according to Owen (1998), has wider implications in so far as the value of something is assessed. Findings, regarding the score obtained, are part of evaluation according to Erasmus (1997) and Owen (1998). They therefore agree that it is better to refer to assessments or measures of cognitive abilities, rather than of tests or evaluations.

De Beer (2000b, p. 1), the developer of the other psychometric test used in this study, namely the LPCAT, pointed out that:

When used properly, tests can provide broader and more equitable access to education, training and employment with individuals, institutions and society benefiting when testing helps to afford individuals opportunities to achieve their goals.

Psychometric instruments are used to obtain information about individuals or groups. Tuckman (1975) indicated the following uses of psychometric tests:

- To give objectivity to observations
- To elicit behaviour in controlled conditions
- To anticipate the unseen
- To detect characteristics of behaviour
- To predict future behaviour
- To make data available for decision making

These uses of psychometric tests make them useful as tools for student selection purposes, specifically the uses of anticipating the unseen and predicting future behaviour.

2.5.1 The basic requirements for psychometric tests

The basic requirement for scientific testing is that all psychometric tests should be reliable, valid and have standardised measures or norms. The extent to which the psychometric instruments used in this study, the PIB and LPCAT, meet these requirements, will be investigated and discussed in chapter four.

2.5.1.1 Reliability

Reliability refers to whether the test results would be consistent over time (Anastasi, 1988; Baron, 1994; Cronbach, 1990; Hoy & Gregg, 1993). In other words, if you were to test the same person twice, reliability refers to whether you would get the same or similar test results with each test (Anastasi, 1988; Baron, 1994; Cronbach, 1990; Hoy & Gregg, 1993; Tuckman, 1975).

Hoy and Gregg (1993) refer to three types of reliability, namely test-retest reliability, internal consistency reliability or split-half reliability and inter-rater reliability. They will now be described briefly:

a) Test-retest reliability

This indicates the stability or ability to generalise over time. This is determined by administering a test to a group on two different occasions. The scores of the test and retest should be strongly correlated (Hoy & Gregg, 1993; Tredoux & Durrheim, 2002). Test-retest reliability is measured by calculating the correlation coefficient between the two scores of the two administrations of the test (Tredoux & Durrheim, 2002).

b) *Internal consistency reliability or coefficient alpha*

This is an estimate of the consistency of responses to different test items. It is “the average of the reliability coefficients that would result if all possible split-half analyses were performed” (Tredoux & Durrheim, 2002, p. 213). Internal consistency leads to test-retest reliability. It is strongly affected by the number of items in the test.

c) *Split-half reliability*

Split-half reliability is when the items of the test are divided into two equivalent halves, in terms of difficulty and coverage, and then the whole test is administered on a single occasion. The scores of the two halves are then correlated to determine whether they provide similar measures. The rationale is that, if the scores are strongly correlated, it implies that if the whole test is administered on two separate occasions, it will also have test scores that strongly correlate. A problem with split-half reliability is that the way in which the test is divided influences the result (Tredoux & Durrheim, 2002).

d) *Inter-rater reliability*

This type of reliability considers the similarity between observer ratings, expressed as a percentage.

e) *Parallel form reliability*

Requires two test administrations where a reliability coefficient is calculated on the scores of the two measures taken by the same group of subjects. A high and positive correlation is necessary to say the two forms are parallel (Heffner, 2004).

Information regarding the reliability of a test is usually included in the test manual or technical manual of psychometric tests (Cronbach, 1990; Hoy & Gregg, 1993). It is essential to follow the standardised procedures which were used on the norming sample to ensure reliability. The reliability coefficient and the standard deviation are used to

calculate the standard error of measurement. This helps to determine the technical quality of a test (Anastasi, 1988; Hoy & Gregg, 1993).

2.5.1.2 Validity

Validity refers to whether a test measures what it claims to measure (Cronbach, 1990; Tuckman, 1975). This is a vital attribute of a test (Baron, 1994; Cronbach, 1990; Hoy & Gregg, 1993). In the case of intelligence tests, the question will be whether they really measure cognitive ability. Anastasi (1988, p. 139) indicates that no test “can be said to have high or low validity in the abstract. Its validity must be established with reference to the particular use for which the test is being considered”. The question of test validity is controversial, as it also addresses the issue of whether a test is fair to use for groups from different cultures (Baron, 1994). This will be discussed in more detail in section 2.6.2.

The validity of a test is determined by calculating the correlation between scores obtained in the measuring instruments and an independent objective measure of performance (Van Aswegen, 1997). Correlation and factor analytical techniques are used to calculate validity (Baron, 1994).

Different kinds of validity exist, namely face validity, content validity, construct validity and criterion-related validity. They will be discussed below.

a) Face validity

This refers to whether the test appears to measure what the test title indicates. This is not a statistical judgment and does not necessarily imply content validity (Anastasi, 1988; Anastasi & Urbina, 1997; Hoy & Gregg, 1993).

b) Content validity

To determine whether a test is valid we need to consider the content. This is another non-statistical judgment considering whether the items included in the test cover the field measured, in other words, whether the content of the test is closely related to the behaviour that one is trying to predict (Anastasi, 1988; Baron, 1994; Hoy & Gregg, 1993; Van Aswegen, 1997). A content valid test will almost always demonstrate ‘face’ validity, but a test with face validity does not necessarily have content-orientated validity. Panels of experts usually judge the validity of the content of a test.

c) Construct validity

According to Hoy and Gregg (1993, p. 73), construct validity “relates to how meaningfully a test measures a psychological trait”. The basic requirement of this kind of validity is an understanding of the trait being measured. A theoretical model of the specific trait is required to determine the construct validity of the test. There are two aspects that are considered, namely:

- the psychological *construct being measured*, and
- *how well* the test measures the construct.

Construct validity is assessed through factor analysis and correlation investigations, comparing the degree of correlation between a certain test and other tests measuring the same construct.

d) Criterion-related validity

Criterion-related validity determines whether the predictor is associated or correlates with the criterion. It is calculated by measuring how close the test scores match the criterion in terms of present or future performance. This is determined statistically by means of correlation (Anastasi, 1988; Hoy & Gregg, 1993).

There are two types of criterion-related validity, namely concurrent validity and predictive validity.

- Concurrent criterion-related validity is used when predictor and criterion scores are obtained at the same time (Baron, 1994).
- Predictive validity is orientated towards the future and is determined by using test scores to predict future performance (Hoy & Gregg, 1993). Van Aswegen (1997, p. 20) indicated that “it involves a time interval between the collection of data on the measurement procedures (predictor) and the performance domain (criterion)”. The extent to which a test successfully predicts future performance, determines the predictive validity of a test.

A predictive validity investigation is the focus of this research project. The aim is to determine the predictive validity of two psychometric tests and matric results as predictors of future academic performance of students over a three-year period. The criterion will be the academic performance of students. The actual performance, in terms of percentages obtained and the number of subjects passed over the three-year minimum period, will be used as criterion.

2.5.1.3 Standardisation and norms

“Standardisation implies *uniformity of procedures* in administering and scoring of tests” (Anastasi, 1988, p. 25). In the development of a test, test constructors should provide detailed instructions regarding the administering of the test. Specific information regarding the materials to be used, time limitations, oral instructions, demonstrations and handling enquiries are some of the detail which should be provided in an instructor’s guide accompanying the test material. Testing therefore should take place in controlled conditions, as with other scientific observations (Anastasi, 1988). Standardised

procedures need to be followed to ensure reliability. If standardised procedures are not followed, it makes the test results unreliable.

The establishment of norms is a component of the standardisation of a test. Anastasi (1988, p. 26) indicated “a norm is the normal or average performance”. Psychological test scores do not represent a pass or fail. Norms provide a way to evaluate the performance of an individual and make the interpretation of test scores possible (Anastasi, 1988; Tuckman, 1975). This enables the test user to compare an individual’s test score with the norms provided. To standardise a test, it is applied to a large group of people for whom the test is designed. This is called the standardisation sample and is used to establish norms (Anastasi, 1988).

Tuckman (1975) refers to two shortcomings of norms:

- People and cultures change over time. If norms are not adjusted it could result in misinterpretations.
- The relative interpretation of tests can obscure the relation between the content of tests and bearing on the past or future reality of test takers.

As norms provide a basis for the interpretation of individual scores, attention will be paid to the standardisation or norm groups of the two psychometric tests used in this project, to ensure that they are fair for the sample.

Specific attention will be given to the standardisation process of the two psychometric tests used in this project. This will be discussed in section 4.4.2.2 and section 4.4.3.2.

2.6 CROSS-CULTURAL MEASUREMENT WITH SPECIFIC EMPHASIS ON THE SOUTH AFRICAN CONTEXT

As reflected in the discussion regarding the requirements of psychometric testing, it is essential to consider cultural differences, throughout test development and administration to the use of norms and the interpretation of test scores.

Cross-cultural assessment is viewed as a complex problem in South Africa (Retief, 1988). There is general resistance against the use of psychometric tests, because of the impact that apartheid policies had on test development. Despite similar resistance against psychometric measurement in the United States of America in the 1970's, psychometric testing survived internationally (Foxcroft, 1997; Owen, 1998).

Biesheuvel, who was involved in the establishment of the National Institute for Personnel Research (NIPR), made major contributions to the literature on cross-cultural research in South Africa between 1940 and 1960 (Biesheuvel, 1943, 1973; Retief, 1988). According to Retief (1988), Biesheuvel pointed out that, in order to test abilities in a culture, a sound understanding of the skills and the knowledge of the target population is necessary for setting culturally appropriate items. Biesheuvel (1973) investigated the issue of culture-free and culture-fair testing in South Africa. He indicated that culture fairness should not be a concern, but that the focus should rather be on the measurement of individual differences. He further stated that there should be a focus on tests studying the modifiability of behaviour, which he referred to as tests of adaptability. This view supports the modern approach of dynamic assessment. According to Biesheuvel (1973) and Jensen (1980), intelligence tests, or tests of mental ability, differ from tests of adaptability. Traditional intelligence tests are regarded as static in nature and only measure learnt skills, while tests of adaptability measure what people can learn to do. Traditional tests are unsuitable to use, due to the differences in educational background of the South African population. Adaptability tests include a learning component. According to Jensen (1980), there is a belief that the only true ability is the ability to learn and that IQ tests only measure what the individual has already learnt. Van Aswegen (1997) pointed out that the popularity of learning potential tests in South Africa is a result of the promise of these tests to provide a fair measure for everybody, despite unequal and inadequate educational opportunities.

2.6.1 Major problems in cross-cultural assessment

Retief (1988) highlighted the major problems in cross-cultural assessment as comparability, translation of meaning and test bias. These aspects will now be discussed.

2.6.1.1 Comparability

Feuerstein et al. (1987) indicated that individuals from culturally different groups could not reasonably be compared with individuals from groups for which the test was designed.

If test scores are not comparable across groups, the possibility of unfair treatment of a particular group increases. This can result in unfairly awarding scarce opportunities to certain groups (Taylor, 1994). This concern results in considering new measures, like dynamic assessment, which may be more in line with a process of redressing imbalances of the past (Taylor, 1994).

2.6.1.2 Translation

Hoy and Gregg (1993) pointed out that there is a risk of inaccurate assessment of true ability when an individual, with limited English proficiency, is subjected to an English test. This can be a potential source of test bias against black pupils in South Africa who completed a test in English (Owen, 1998).

De Beer (2000c, p. 6) pointed out that:

When a person who is not fluent in a language is tested in that language, the resulting score may well be more an indication of the person's language proficiency, than of the skill that is supposedly measured.

Another concern is that when a test is used across cultures, the tester needs to communicate across cultures. Communication refers to the process through which

information (meaning) is transmitted from the sender to the receiver. The meaning of a message can be distorted due to misunderstanding or misinterpretation, if it is not in the mother tongue of the testee. If tests are translated into black languages, much of the content can lose its meaning. Although high quality translations are available, psychological constructs vary between cultures (Van Aswegen, 1997). Simple translation is therefore not sufficient and some adaptation and revision is required (Anastasi, 1988.) It is therefore often difficult to validate tests for black samples (Van Aswegen, 1997).

Non-verbal content is sometimes used in an attempt to obtain more culture fair measures (Anastasi, 1988). Although some tests have figural content and therefore depend less on communication, they always include language, even if it is only the instructions for the non-verbal procedures (Retief, 1988).

To increase the validity of a test by preventing the language or translation problem, it should rather be redeveloped in terms of the meaning system of a culture (Retief, 1988).

2.6.1.3 Test bias

Test bias is the last problem of cross-cultural assessment that Retief (1988) refers to. In the literature many different definitions for test bias exist. The technical or psychometric understanding of bias refers to whether the test is biased against a disadvantaged group, meaning that the test discriminates against a group, or is unfair towards a particular group. Bias can also be understood as the validity of the test or test items for a certain group (Huysamen & Raubenheimer, 1999; Jensen, 1980; Retief, 1988).

Jensen (1980) views test bias as a statistical concept that refers to systematic errors in predictive and construct validity that are associated with group membership. He does not believe that tests measure different constructs in different cultures. Cronbach (1973) indicated that regression analysis could be done if one needs to defend the use of psychometric instruments in a selection procedure on different cultural groups. Different

kinds of test bias exist, namely construct, method and item (or differential) bias (Van de Vijver & Poortinga, 1997; Van de Vijver & Tanzer, 1997). These different forms of test bias are described below.

a) Construct bias

Construct bias concerns the nature and essence of what is being measured (Owen, 1998). Construct bias appears when “the construct being measured is not identical across cultural groups” (Van de Vijver & Tanzer, 1997, p. 264). This usually happens when test designers come from different cultures or societies and define the construct being measured differently (Van de Vijver & Poortinga, 1997).

The following are sources that can result in construct bias (Van de Vijver & Tanzer, 1997):

- Little overlap in the definition of the construct between cultures.
- Appropriateness of behaviour associated with the construct in different cultures.
- Poor sampling of behaviour associated with the construct.
- Aspects of the construct are not completely covered.

Figural content can be used in tests to reduce the influence of the culture on answering the test items, but even figures can be culturally entrenched (Van de Vijver & Poortinga, 1997).

b) Method bias

Method bias occurs when a cultural factor, not relevant to the construct being measured, influences most of the test items, for instance communication or language proficiency. In the administration of a test where the first language of tester and testees differs, the collection of data can be influenced (Van de Vijver & Tanzer, 1997).

Method bias can have serious consequences on the validity of cross-cultural comparisons, as the average scores of groups or cultures can differ (Van de Vijver & Tanzer, 1997).

Possible sources of method bias, as described by Van de Vijver and Tanzer (1997), are:

- Samples not comparable, because of differences in background, like education.
- Differences between the environments where the tests are administered.
- Ambiguous instructions for respondents and/or guidelines for administrators.
- Different levels of expertise of test administrators.
- The halo effect in tester/viewer/observer.
- Communication and interpretation problems.
- Differences in familiarity with test material.
- Differences in familiarity with response procedures.
- Differences in response styles.

c) *Item bias*

This kind of bias is also called differential item functioning (Van de Vijver & Poortinga, 1997; Van de Vijver & Tanzer, 1997). Van de Vijver and Poortinga (1997, p. 30) report that an “item is biased if persons from different groups with the same score on the construct ... do not have the same expected score on the item”.

Jensen (1980), Owen (1998) and Retief (1988) indicate that item bias occurs when items, unknown to a certain culture, are included in a test. Cronbach (1990) points out that, when the content is reviewed, each item should be judged in terms of whether it measures what the tester wants it to measure. According to Eysenck (1971), tests may contain material that is learned and may be peculiar to a particular culture and it might be more appropriate to call them tests that measure education rather than intelligence.

Standardised tests are often loaded with items based on middle-class values and experiences, because the tests are developed by middle-class people and standardised for a middle-class population (Babad & Budoff, 1974; Hegarty, 1988; Jensen, 1969). This can result in measures being unfair to individuals from lower socio-economic levels in South Africa's multicultural society.

Statistical procedures, for example the Item Response Theory (IRT) (Claassen, 1990a), ANOVA (Analysis of Variance) and Mantel-Haenszel statistics, can be used to detect biased items (Van de Vijver & Poortinga, 1987, 1997). De Beer (2000c, p. 38) indicated "whichever technique used, a crucial issue is the clarity and theoretical as well as practical justification of the definition of the groups that will be compared".

Van de Vijver and Poortinga (1997) summarise the sources resulting in item bias as:

- Poor translation of test items.
- Inadequate item formulation.
- Items invoking additional traits or abilities.
- Incidental differences in the appropriateness of item content.

d) *Predictive bias*

This is also called differential prediction (Huysamen & Raubenheimer, 1999, p. 172), which "is concerned with whether there is bias present in a particular predictor or prediction system". This is a statistical or psychometric concept, which means that the criteria scores used are systematically underpredicted or overpredicted for members of a certain demographic group. This can be illustrated by considering a situation where applicants from a particular demographic group consistently perform better or poorer than what was predicted. If this happens in the case of educationally disadvantaged students, it can have serious political consequences (Huysamen & Raubenheimer, 1999).

Differences in the intercepts and/or slopes of the criterion-on-test regression lines of different demographic groups signal the presence of differential or prediction bias (Huysamen & Raubenheimer, 1999, p. 172).

To overcome the concerns raised, attention should, firstly, be paid to the construction of tests. Secondly, direct comparisons of scores should be avoided and thirdly, bias can be dealt with by identifying the effects statistically and correcting the test to meet the psychometric standards (Poortinga, 1995).

2.6.2 The issue of culture fairness

The fairness of psychometric tests is an issue both locally and internationally (Jensen, 1980; Zeidner, 1988). Fairness in testing is under the spotlight in South Africa, due to our multicultural society and the need to address imbalances of the past. Fairness in psychometric testing used for student selection is essential, as it has implications that go beyond education and psychology, touching on political issues, because measures of cognitive abilities are used to provide scarce opportunities (Shochet, 1986).

Fairness can be distinguished from bias, as it concerns the use of the test scores, while bias concerns the influences on scores during testing. The aim of culture-fair testing is to reduce or eliminate effects of cultural differences on scores, as test scores that do not represent an equal measurement can lead to unfair decisions (Retief, 1988). Fair test use does not require culture fairness of test instruments (Poortinga, 1995).

According to Jensen (1980), tests are culturally unfair if they discriminate against racial groups or persons of low socio-economic status. This can also happen when a test, that was standardised for a certain population group, is used on a sample that is not the same as the population group for which it was standardised (Gredler, 1988; Jensen, 1980).

Requirements of tests to ensure culture fairness, according to Eysenck (1971), are that tests should measure:

- Differences between abilities within one person.
- Differences between individuals with regard to a particular ability.
- Differences between the averages of groups.

Anastasi and Urbina (1997, p. 164) stated the following regarding culture fairness of tests:

If we want to use tests *to predict outcomes in some future situation*, such as an applicant's performance in college or job, *we need tests with a high predictive validity* against the particular criterion. This is often overlooked in the development of so called culture-fair tests. In the effort to include in these tests only functions common to different cultures or subcultures, *we may choose content that has little or no relevance to any criterion we wish to predict*. [italics added]

A test will be regarded as fair and valid if the predictor variable has a positive relation to the criterion (Poortinga, 1995; Retief, 1988).

These issues, and the emphasis of the Department of Education (1997), as well as the Employment Equity Act, No. 55 (1998), indicate the need for fair, reliable and valid instruments that can be used to award scarce opportunities. Foxcroft (1997) and Retief (1988) pointed out that the development of fair testing practices and assessment was slow in South Africa during the 1990's, because it was more complex than what was initially thought.

Although cultural fairness and cross-cultural validity are not investigated in this study, the literature review enables the researcher to report on the apparent cultural fairness of

the psychometric tests used in the study. Both tests used in this project aimed at being culture-fair tests. Each of the two tests was developed in South Africa and was standardised for South Africa's multicultural population. Both tests also use figural content to overcome possible language and/or communication problems.

2.6.3 The future of psychometric testing in South Africa

The problems and concerns regarding psychometric testing in South Africa, discussed in section 2.6.1, as well as the legal requirements of the Employment Equity Act, No. 55 (1998), necessitated certain changes in psychometric tests in South Africa.

According to Van Aswegen (1997), it is generally agreed that psychometric assessment in South Africa was at a crossroad in the late 1990's. With the socio-political changes in the late 1980's, discontinued job reservation and mixed schools resulted in industry and educational authorities demanding common tests that would not unfairly discriminate against anyone, irrespective of race or culture (Claassen, 1995).

Foxcroft (1997) argued that resistance to psychometric tests was beneficial to psychometric testing in South Africa, as it forced practitioners and test developers to reflect on the purpose, and the ethical use of psychological tests in our multicultural society.

Taylor (1994, p. 193) indicated that:

If we are to address the inequalities of the past in South Africa, employers and educationalists will have to place *more emphasis on potential rather than skill or specific ability*... he further added that ...we should be prepared to give those with high potential opportunities to develop.

Foxcroft (1997) indicated that there is an urgent need for the adaptation of current tests or the development of culturally relevant tests and norms. This is essential to enhance psychological testing and assessment practices in South Africa. She indicated that during the apartheid years, when education and jobs were reserved for certain groups, it was almost inevitable that psychological tests also developed along cultural lines. Therefore, the users and developers of psychological tests now face numerous challenges. This will enhance the fair and ethical use of tests.

There were two compelling reasons that hindered the development of tests in the 1990's , according to Foxcroft (1997) and Owen (1998). The first reason was personnel rationalisation and financial constraints at the HSRC. This was further complicated by a lack of clarity regarding the future core focus areas of the HSRC in the new political dispensation. Owen (1998) indicated that the HSRC was viewed with suspicion as a pro-apartheid think-tank. This had a further negative effect on the role they played in test development. The second reason was based on the inherent difficulty of developing tests in a multicultural and multilingual society. Developing tests that focus on behavioural aspects and cognitive functioning that are common to various cultures of South Africa, is highly complex (Foxcroft, 1997). To overcome this, members of all cultures concerned should be actively involved in the development of culturally common tests (Taylor & Boeyens, 1991).

Huysamen (1996b), Foxcroft (1997) and Van Aswegen (1997) emphasised that psychometric assessments should never be used in isolation. The use of demographic and other data, in combination with psychological test results, could result in a fair and unbiased admission procedure at tertiary institutions.

Both psychometric tests used in this project represent the changes required, like focusing on the potential of testees rather than measuring crystallised abilities. The use of such tests in student selection can contribute towards addressing the imbalances in South Africa through tertiary education. The sample group for the present study was from the same cultural group and consequently cross-cultural comparison was not possible.

2.7 APPROACHES TO COGNITIVE ASSESSMENT

Various authors distinguish between three main approaches in the measurement of cognitive abilities (Cronbach, 1990; Gardner, 1993; Hoy & Gregg, 1993; Owen, 1998; Taylor, 1994). These approaches are the conventional or structural approach, the information processing approach and the learning potential or dynamic testing approach. According to Taylor (1994), these approaches developed fairly independently from one another. Other sources of literature also refer to the Piaget approach (Cronbach, 1990; Gardner, 1993; Hoy & Gregg, 1993; Owen, 1998).

2.7.1 The conventional or structural approach

Tests developed under this approach are based on psychometric theories of intelligence (Sternberg, 1985). This approach measures individual differences in performance along dimensions that represent the fundamental structure of the psychological domain. Domains that are included are cognition, personality, interests and others. Individual differences of the cognitive domain are measured with a power score, which is the adding of correct answers to problems. The speed at which problems are answered is not regarded as important in this approach (Hoy & Gregg, 1993; Taylor, 1994). An essential part of this approach is the adherence to standardised procedures in the administering of the test (Hoy & Gregg, 1993). Correlation or factor analysis is used to resolve theoretical and empirical questions (Owen, 1998; Sternberg, 1985; Taylor, 1994). Tests from this approach are static and provide no feedback to the testee during testing (Grigorenko & Sternberg, 1998; Owen, 1998; Sternberg, 1985, 1997)

There are two schools of thought that resort under this approach. The one school of thought defines intelligence as a general unified capacity (Hamers & Resing, 1993). This approach started with Galton's laboratory and reaction time studies in London over a century ago (Taylor, 1994). The theories of Spearman and Cattell also resort under this approach (Sternberg, 1997; Taylor, 1994). The second school of thought assumes that

intelligence consists of separate mental abilities as formulated in the theories of Thurstone, Guilford and Gardner (Hamers & Resing, 1993).

The main criticism against this approach is the basic assumption that all people have equal opportunities to obtain the knowledge and skills that a test requires. Another problem with this approach is that the content is loaded with items dominant in a specific culture (Van Aswegen, 1997). Hoy and Gregg (1993), as well as Taylor (1994), indicated that, because IQ tests are used to award scarce opportunities such as admission to tertiary institutions and jobs, those who do not qualify might feel negative towards these tests as they see them as denying them opportunities. Taylor (1994) argued that if these tests were still to be used in a context of affirmative action, new ways of using the test scores would have to be considered, for example using separate norms.

Owen (1998) agrees with Taylor (1994) that this approach has limitations for use with culturally different groups, unless norms are standardised for the different cultures.

The PIB, one of the psychometric tests used in this project, resorts under this approach. The developers did, however, attempt to overcome the concerns regarding cultural fairness expressed against this approach. The PIB uses non-verbal test items to overcome language problems. The developers of the test also reported that items were selected carefully to prevent the inclusion of biased items. Attention was also paid to the concern that Owen and Taylor expressed regarding test norms and standardisation, and the PIB was standardised using a sample that is representative of the South African population. This test will be discussed in more detail in section 4.4.2.

2.7.2 The information processing approach

This approach is also known as the cognitive theories of intelligence (Sternberg, 1985), and was established in the 1960's as a basis of neuropsychology or cognitive psychology (Hamers & Resing, 1993; Taylor, 1994). The central component of this approach is the

way in which humans *process* information. Two traditions developed within the information processing approach. The cognitive components approach of Sternberg, which aims to uncover components of cognitive thought and the cognitive correlate approach which focuses on the step-by-step analysis of cognitive processes that support intelligent behaviour (Cronbach, 1990; Hoy & Gregg, 1993; Hamers & Resing, 1993; Sternberg, 1985). This information processing approach intends to find critical processes that underlie performance, for example verbal ability (Taylor, 1994).

None of the tests used in this project resorts under this approach.

2.7.3 The Piagetian approach

In addition to the three other approaches that Taylor (1994) refers to, Cronbach (1990), Gardner (1993), Hoy and Gregg (1993) and Owen (1998) also mention the Piagetian approach. As indicated in section 2.3, the Piagetian approach evaluates the development of individuals at four different stages (Owen, 1998; Sternberg, 1985). This approach is useful to assess pre-school children. The primary focus of the Piagetian approach is to investigate reasoning abilities (Hoy & Gregg, 1993). This approach differs from the psychometric approach in that it does not focus on individual differences, but rather on the development of cognitive abilities of individuals (Owen, 1998).

This approach was not used in the present study.

2.7.4 The learning potential or dynamic testing approach

This approach is also referred to as the Soviet approach, because the theory that supports this approach most is that of the Russian psychologist, Vygotsky (Hoy & Gregg, 1993; Owen, 1998; Vygotsky, 1978). The learning potential or dynamic testing approach is viewed as a viable way to address psychometric testing or assessment in culturally

different and disadvantaged populations. According to supporters of the dynamic assessment approach, one of the concerns regarding conventional ability testing is the underestimation of real intellectual potential (Hamers & Resing, 1993; Owen, 1998). Feuerstein (1973) and Shochet (1986) point out that traditional tests are unable to distinguish between manifest and potential performance, and therefore support the view of Owen (1998) that dynamic assessment should be considered.

This approach views intelligence as the human capacity to adapt to demands of the environment. The dynamic approach focuses on the assessment of adaptation (increased speed and accuracy) as a result of repeated exposure or instruction (Taylor, 1994). Vygotsky viewed cognitive competence as a social phenomenon and defined the concept “zone of proximal development” as acknowledgement of individuals’ different capacities to benefit from mediating learning experiences (Hoy & Gregg, 1993; Owen, 1998). The aim of the dynamic assessment approach is to distinguish between unassisted performance and performance attained through mediation or after some form of training or learning (Taylor, 1994; Vygotsky, 1978).

In dynamic assessment the testee engages in a learning process, while the tester attempts to facilitate cognitive competence (Owen, 1998). The aim is to determine how the testee profits from the assistance (Budoff, 1987a; Owen, 1998). The administering of a dynamic test involves a training component (Cronbach, 1990; Hoy & Gregg, 1993; Owen, 1998).

Dynamic testing consists of two main approaches. The enrichment approach aims at the remedial value of dynamic assessment. The best-known example of this approach is Feuerstein’s Learning Potential Assessment Device (LPAD) (Owen, 1998). There is also the measurement approach that focuses on the change in performance, such as the improvement between pretest and post-test scores that Vygotsky refers to as the “zone of proximal development” (ZPD) (Das, 1987; Owen, 1998). According to De Beer (2000c), whose learning potential test resorts under this approach, learning potential is a

combination of the actual (initial pretest score) and the ZPD or difference score. These two approaches will be discussed in more detail in section 2.8 and 2.9.

Other approaches that have been designed to assess learning potential are, amongst others, that of Budoff as well as Campione and Brown (Sternberg, 1997; Taylor, 1994). These tests will be discussed in section 2.9.3.

One of the other psychometric tests used in this project, namely the LPCAT, falls under the measurement approach of learning potential approach. It uses a test-train-retest approach with standardised training built into the test. This test will be discussed in detail in section 4.4.3.

According to Hamers and Resing (1993), the information processing approach and the learning potential approach differ in their aims. The aim of the information processing approach is to measure cognitive processes, while the learning potential approach aims to train an individual, and then determines whether the individual benefited from the training and assesses the efficiency of the cognitive processes.

2.8 TRADITIONAL INTELLIGENCE TESTS VERSUS DYNAMIC ASSESSMENT

According to Hamers and Resing (1993), cognitive abilities can be measured in two different ways, namely as an IQ score or as a learning potential score. This refers to traditional psychometric tests and dynamic tests. Traditional views on cognitive abilities were discussed in section 2.3 and dynamic assessment instruments will be discussed in section 2.9.

Campione and Brown (1987) raised the question as to the differences between static and dynamic assessment. They concluded that static tests, also called traditional tests, evaluate the current ability resulting from cultural and educational background, while

dynamic assessment assesses learning potential. There is some criticism against conventional IQ tests regarding construct validity. The concern is that they focus on products rather than on cognitive processes and therefore produce no information on learning potential or cognitive modifiability (Hamers & Sijtsma, 1993; Lidz, 1987). According to Guthke (1992), the testing of current abilities is even more inadequate now than in the past, because we live in an era of lifelong learning and rapid technological change. What is more important is the ability of individuals to adapt to the demands of the environment in which they exist. Jensen (1969), as well as Sternberg (1985) sixteen years later, indicated that there was a need for other standard tests, measuring a broader spectrum of mental abilities than traditional tests. Jensen (1969) referred to this broader spectrum of mental abilities as learning ability. According to Guthke and Stein (1996), dynamic tests assess acquisition of knowledge better than conventional tests.

The administration of static, traditional and dynamic measuring instruments also differ. In dynamic assessment, other than in traditional tests, there is some form of intervention, usually coaching or training, during the testing process. Dynamic assessment is based on the theoretical notion that intelligence is a dynamic entity and it assesses the possible change in ability resulting from training (Feuerstein et al., 1987; Grigorenko & Sternberg, 1998; Gupta & Coxhead, 1988; Sewell, 1987). Learning potential tests reveal that individuals differ in taking advantage of training. The information regarding these differences contributes towards prediction of future learning (Ruijssenaars, Castelijns & Hamers, 1993). Budoff (1987b) therefore argues that dynamic assessment provides an opportunity for individuals to demonstrate how they can perform when they understand the demands of the task. The assessment of learning potential prevents underestimation of the potential of individuals from disadvantaged backgrounds, which makes these tests very suitable to use in student selection. Budoff (1987b), Gupta and Coxhead (1988), as well as Sternberg (1985), indicated that, according to proponents of the dynamic approach, training-based assessment is better than measures of current levels of functioning, because individuals come from different cultures and backgrounds with different opportunities to acquire the ability being measured.

In the two approaches of dynamic assessment the form of intervention differs. In the enrichment approach of dynamic assessment coaching or training is used, with the mediator aiming to facilitate changes in the cognitive functioning of an individual. This approach requires a lot of time and effort from the administrator, which makes it very expensive (De Beer, 2000c; Hamers & Resing, 1993; Lidz, 1987). The measurement or psychometric approach of dynamic assessment uses a pretest-training-post-test procedure. The focus is on the standardisation of training and assessment to increase the accuracy in measurement to obtain an objective, valid, reliable and quantifiable measure of learning potential. Three scores are obtained, namely a pretest score, a post-test score and a difference score.

Budoff (1987b) and Guthke (1993) stated that despite constant debate regarding the shortcomings of IQ tests, they are still in use. Professionals and clients still prefer to use traditional methods, knowing that they are inappropriate, rather than use the practice of assessment to guide the educational process (Feuerstein et al., 1987). Feuerstein et al. (1987) further argued that conventional methods should either be abolished, or their structure should change drastically. In Israel they went as far as prohibiting the use of static measures in schools (Brown & French, 1979; Feuerstein et al., 1987). An alternative to the abolishment of these tests is that they should be substituted by different techniques.

The measurement of learning potential is therefore an alternative strategy to assess cognitive ability (Ruijsenaars et al., 1993). With this method the assessment is not restricted to current ability as with traditional static tests (Budoff, 1987b), and it is further indicated that this method of assessment permits reasonable predictions regarding educability. The learning potential paradigm places all individuals on a more equal base, as everyone is trained in relevant problem-solving strategies (Babad & Budoff, 1974).

According to Taylor (1994), the focus of testing in South Africa today should rather be on testing for potential or learning potential to address the inequality of opportunities in the South African society. Taylor indicated that this would contribute to the process of

reconstruction and development. Foxcroft (1997), Owen (1998) and Retief (1988) supported Taylor and indicated that it would be wise to use tests that measure potential, rather than those reflecting on educational background. The two psychometric instruments used in this study have overcome the concerns expressed about traditional measures. The aim is to assess their suitability in predicting future academic performance.

2.9 THE DEVELOPMENT OF DYNAMIC ASSESSMENT APPROACHES AND INSTRUMENTS

Dynamic assessment approaches developed from learning, developmental and cross-cultural psychology (Hamers & Resing, 1993). The first tests in the dynamic assessment or learning potential approach started as the repackaging of existing IQ tests (Hamers & Resing, 1993). Dynamic tests are also sometimes referred to as long-term learning tests, because of the pretest, training phase and post-test design. The dynamic assessment methods use a number of different techniques that involve different forms of assistance to individuals (Campione & Brown, 1987). The methods also differ in terms of their assessment methods and their goals. There are mainly two approaches to dynamic assessment, namely the enrichment approach and the assessment or measurement approach (Minick, 1987).

In the clinical or enrichment approach, capacity is viewed as modifiable as a result of cognitive modification. In view of this approach structural changes of cognitive ability take place. This approach makes use of mediation techniques and measures changes qualitatively (De Beer, 2000c; Grigorenko & Sternberg, 1998; Murphy, 2002). The measurement of learning ability falls within the context of experiential research, where time is required for the observation of effects (Lidz, 1987).

The assessment or measurement approach, also known as the psychometric approach, focuses on the quantitative measurement of the changes in performance (Minick, 1987).

Lidz (1987) indicated that activity and modifiability are two keywords of the dynamic assessment concept. Activity refers to the process between the person who intervenes through training and the examinee, with the aim to modify the ability or performance of the individual. This implies that there is interaction between an examiner as intervener and the learner as active participant. The interventions are part of a test-teach-retest process within the measurement or assessment approach. Modifiability refers to the capacity to learn if the opportunity is provided (Grigorenko & Sternberg, 1998; Murphy, 2002). The dynamic assessment measures under this approach focus on the standardisation of test procedures to improve the measurement accuracy and to provide objective, valid, reliable and quantifiable measures of learning potential. The focus is therefore more on psychometric principles of assessment than is the case with the enrichment approaches (De Beer, 2000c).

2.9.1 Aims of dynamic assessment

According to Minick (1987), the dynamic assessment movement recognised the need to develop assessment devices that provide the following:

- Direct measures of potential for learning and development.
- Information on processes that result in success or failure of individuals.
- Information on what can be done to facilitate education and development of individuals.

The aim of the dynamic assessment, according to Hamers & Resing (1993), is to obtain a more realistic measurement of a person's cognitive ability, through the inclusion of training as a part of the assessment process. In this way the extent of individuals' sensitivity to instruction and the efficiency of operation of cognitive processes can be determined.

Embretson (1987) pointed out three aims of learning potential assessment:

- To provide a better estimate of the ability construct.

Differences between testees are overcome through training. The training also helps to eliminate the unfamiliarity of testing. These factors improve the validity of dynamic test scores and increase the comparability of individual differences (Grigorenko & Sternberg, 1998).

- To measure new abilities such as the modifiability of performance.

Readiness for change is determined by measuring the amount of help that is needed by the individual (Embretson, 1987) or the magnitude of improvement after training.

- To improve mental efficiency.

This can be accomplished through extensive training. The level of ability therefore changes (Embretson, 1987; Grigorenko & Sternberg, 1998). The dominant purpose, according to Grigorenko and Sternberg (1998), is modifying, changing and improving cognitive performance.

In the development of a learning potential test in South Africa, De Beer (2000c, p. 135) indicated that the aim “is to avoid material that is related to socioeconomic background and to move away from the measurement of that which has been learnt previously”. The dynamic assessment approach therefore reduces the potentially negative effects of a lack of previous life experiences on test performance. The negative effects are reduced through repeated contact with the material in a supportive and teaching environment (Babad & Budoff, 1974). This is why dynamic assessment is so attractive to use in South Africa with its varying educational opportunities.

2.9.2 The evolution of the learning potential concept

Dynamic assessment originated in the 1920's. Lidz (1987) summarised the highlights of the development of dynamic assessment as follows:

- Between the 1920's and the 1930's, the definitions of intelligence included the ability to learn and the pace of learning.
- The first experimental investigation of the concept of dynamic assessment was done in the 1930's to the 1940's (Grigorenko & Sternberg, 1998).
- In the 1940's, research findings of Woodrow indicated that intelligence and learning ability are not the same.
- In the 1950's, attempts were made to assess the effects of direct teaching on scores. The results of research projects during this period indicated that individuals with a higher initial IQ benefited from practice, while individuals with a lower initial IQ benefited more from coaching.
- In the 1960's, measures of learning, to assess educability, were developed. During that time Jensen also suggested that dynamic assessment could be used for ethnic minority groups.
- The 1970's were characterised by the work of Feuerstein, Budoff, Campione and Brown. Vygotsky's work on the zone of proximal development was translated into English and propagated in the United States of America. This theory will be discussed in more detail later in this chapter.
- During the 1980's, attention was paid to studies validating some basic premises of dynamic assessment.
- In the 1990's, more instruments to measure learning abilities were developed and introduced.

Dynamic assessment is widely accepted and extensively researched today. In South Africa there were various attempts to develop and/or use dynamic assessment instruments

to overcome the problems of psychometric testing, raised in section 2.6.1. There were also various predictive validity investigations to determine the suitability of dynamic assessment measures to be used in selection batteries at tertiary institutions (Boeyens, 1989a, 1989b; De Beer, 2000a, 2000b, 2000c, 2003; Murphy, 2002; Shochet, 1992, 1994; Van Aswegen, 1997; Zolezzi, 1995).

2.9.3 Dynamic assessment theories and instruments

Boeyens (1989a) indicated that Vygotsky, during the seventies, introduced an alternative view of mental development. Boeyens (1989a) and Sutton (1988) referred to Vygotsky as the theoretical icon of dynamic assessment. Vygotsky was a pro-environment social scientist who advocated intellectual growth through social interaction (Campione & Brown, 1987; Das, 1987; Sutton, 1988; Vygotsky, 1978). Vygotsky was a student of Wilhelm Wundt and William James (Cole & Scribner, 1978) and believed that learning could bring about mental development (Boeyens, 1989a). In Vygotsky's view, mental processes and abilities are the result of the acquisition of a socially determined heritage as individuals co-operate with one another (Guthke, 1992; Hoy & Gregg, 1993). The intellectual environment determines intellectual growth in children, according to the Vygotskian view. This view implies that the quality of learning opportunities plays an essential role, as well as the potential to benefit from instruction (Boeyens, 1989a; Budoff, 1987b; Vygotsky, 1978). The integration of instruction is therefore regarded as a common characteristic of dynamic assessment (Hamers & Sijtsma, 1993).

According to Boeyens (1989a, p. 6):

Vygotsky asserts that if two children of equivalent measured ability are given adequate learning opportunities, it is possible, *by taking their learning potential into consideration, to predict which child will achieve greater cognitive development.* [italics added]

Vygotsky (1978, p. 85) also said, “What children can do with the assistance of others might be in some sense even more indicative of their mental development than what they can do alone”.

2.9.3.1 The learning potential concept

Vygotsky established the learning potential concept (Hamers & Resing, 1993). He indicated that two levels of development are applicable to determine a child’s learning potential (Vygotsky, 1978). The first level represents the actual development of the individual, for example the general ability measured by traditional static tests. This represents the problem-solving skills that the individual has at that point. The other developmental level is the potential developmental level, which represents the problem-solving ability that can be reached with assistance. The difference between these two levels he termed “the zone of proximal development”, known as ZPD.

The learning potential score can be obtained through an interval scale (Babad & Budoff, 1974; Sijtsma, 1993). The first score represents the unassisted problem-solving attempt. After some intervention, the second attempt is then reassessed to obtain a second score (Boeyens, 1989a). The intervention represents a form of training. Training makes it possible for the individual from a low socio-economic background to understand how to solve a problem. This helps the individual for whom the contents of problems are strange and who is unaware of appropriate strategies to find solutions (Babad & Budoff, 1974).

As indicated, learning potential assessment includes training in a test-train-retest sequence (Budoff, 1987b; Sijtsma, 1993). The training embedded in the testing process elaborates on what is expected in the test task, and this enables testees to respond more effectively to the test task. The training aims to improve the test performance of the testee (Sijtsma, 1993). Guthke and Stein (1996) indicated that researchers hoped that prompts, that are included in the testing process, would result in a fairer and more valid assessment than static conventional intelligence tests.

The pretest reflects the individual's current abilities, while the post-test training scores represent the individual's scores once conditions have been optimised, through ensuring that all individuals are familiar with the tasks and demands (Babad & Budoff, 1974; Guthke, 1993).

2.9.3.2 Vygotsky's ZPD

Brown and French (1979), Grigorenko and Sternberg (1998), as well as Guthke (1993), refer to the concept of Vygotsky known as the "zone of proximal development" or the "zone of potential development". This is regarded as the key principle of learning potential assessment (Hamers & Resing, 1993). The word "proximal" is the translation of the Russian word "blizhaishei" which mean nearest (Minick, 1987).

In contrast to the Piagetian view, that mental development facilitates learning, Vygotsky believed that learning and instruction result in mental development (Boeyens, 1989a). The zone of proximal development refers to the possibility of intellectual growth of individuals through social learning interactions. These interactions, according to Boeyens (1989a) and Das (1987), refer to the opportunity that a child has for verbal interaction with adults. This modifies not only the content of thinking, but also its structure. Cole and Scribner (1978) indicate that Vygotsky refers to this as qualitative and quantitative changes that take place.

Vygotsky argued that the aid provided through training determined the level of performance that an individual could reach (Campion & Brown, 1987). According to Brown and French (1979), the assessment of the width of an individual's zone of potential development (ZPD) indicates how many clues were necessary to solve the problem. A child therefore has a wide zone of potential development if he/she needs fewer clues to solve problems following the training, in other words, he/she has learned to transfer new knowledge to similar problems. The zone is regarded as narrow if the individual shows little improvement after the instruction provided (Campion & Brown,

1987; Cronbach, 1990). This refers to an individual's readiness to change (Campione & Brown, 1987).

Boeyens (1989a) pointed out that the ZPD score, as measure of learning potential, is sometimes incorrectly interpreted as a predictor of future academic performance. ZPD indicates the individual's receptiveness to intervention or instruction. It is therefore rather the level to which the ZPD can increase the ability, that can serve as a predictor of future academic performance. It is thus not the difference score (ZPD) that is interpreted as learning potential.

2.9.3.3 Views of other theorists on dynamic assessment

Budoff (1987b), another pioneer in the development of dynamic assessment, indicated that biology and culture play a role in the development of intelligence. Budoff agreed with Vygotsky that intelligence is influenced by the culture of the individual, and that a test measures the extent to which an individual has integrated, and can express, what he has learned in his culture (Budoff, 1974, 1987b; Grigorenko & Sternberg, 1998). Budoff (1987b) indicated that the construct of intelligence cannot be considered independently from the environment, and therefore the measurement of general ability is incorrect. Budoff (1987b) and Das (1987) are of the opinion that assessment alone has little value, unless it can guide intervention.

According to Grigorenko and Sternberg (1998), Budoff's work was based on the assumption that some educable disadvantaged children are more capable of learning than conventional tests suggest.

Budoff and his associates used static instruments in a dynamic manner and made valuable inputs to the standardisation of instructions. According to them, tests should identify individuals who can profit from learning experiences. One of the tests that they administer in a dynamic manner is the Raven's Progressive Matrices (Hamers & Resing, 1993). With the Raven test-practice-retest paradigm, three types of data are obtained,

namely, pre-training scores reflecting the current abilities, post-training scores that represent the optimal level of functioning of the child and the gain scores from pretest to post-test that illustrate the child's ability to profit from instruction or training (Budoff, 1987b; De Beer, 2000c; Gredler, 1988; Grigorenko & Sternberg, 1998). Budoff's approach therefore resorts under the assessment or measurement approach of dynamic assessment.

For Budoff the pretest scores are similar to an IQ test measuring crystallised intelligence, while the post-test scores indicate the optimum ability of an individual (Budoff, 1974, 1987b). Babad and Budoff (1974) describe learning potential as trainability, that is, the ability to improve performance after a learning experience.

Grigorenko and Sternberg (1998) summarised the distinct characteristics of Budoff's approach as follows:

- It is designed as an alternative to conventional tests to select children for special education.
- Only standardised, reliable and extensively validated tests are used and permitted.
- The training component is aimed at familiarising students with demands and equalising their experiences.

Feuerstein also made major contributions to the field of dynamic assessment. Feuerstein's approach resorts under the enrichment approach of dynamic assessment, focusing on structural cognitive modifiability (Feuerstein, Rand & Rynders, 1988; Jensen & Feuerstein, 1987). Feuerstein does not view intelligence as static, and defines it as the capacity to learn from direct exposure to stimuli (Feuerstein, 1980). Learning potential and structural cognitive modifiability of individuals are therefore essential concepts in his view.

Hamers and Resing (1993) indicated that Feuerstein distinguishes between two types of learning, namely direct exposure to a stimulus and mediated learning, where another person mediates the stimulus. The availability of adults to mediate stimulus for children is essential according to Feuerstein. The more mediated learning experiences (MLE), the more the child profits from direct learning experiences and a lack of mediated learning experiences result in deficient cognitive functioning (Cronbach, 1990; Hamers & Resing, 1993). This theory contends that MLE can improve cognitive abilities at any stage of an individual's development (Feuerstein, 1980). MLE refers to the way in which environmental stimuli are transformed by mediating agents (Jensen & Feuerstein, 1987; Shochet, 1986; Tzuriel, Samuels & Feuerstein, 1988). The principles of mediation involve familiarisation with materials, instruction of rules and problem-solving strategies and extensive feedback (Hamers & Sijtsma, 1993). Feuerstein linked cultural deprivation and the issue of mediation. He viewed reduced cognitive modifiability as a result of a lack of exposure to sources of stimulation (Feuerstein, 1980).

Feuerstein and his colleagues developed a test to assess learning potential, the Learning Potential Assessment Device (LPAD) (De Beer, 2000c; Büchel & Scharnhorst, 1993; Das, 1987; Grigorenko & Sternberg, 1998; Hamers & Resing, 1993; Hoy & Gregg, 1993). The LPAD was developed from Feuerstein's extensive research on disadvantaged and retarded children (Cronbach, 1990; Hamers & Resing, 1993; Sewell, 1987). The device is committed to strategies to improve cognitive functioning on a permanent basis. The dynamic assessment model emphasises principles in understanding the process of thinking and behaviour, rather than assesses levels of functioning. The assumption that guided the development of the LPAD was one of individuals as open systems, who are receptive to influences that can produce structural changes in cognitive functioning. It attempts to determine the differences between individuals' abilities to benefit from formal and informal learning opportunities. This structural aspect of change distinguishes LPAD from other dynamic assessment approaches (Feuerstein et al., 1988; Jensen & Feuerstein, 1987).

Feuerstein et al. (1988, p. 198) indicated that there are four major areas where the LPAD differs from conventional approaches, and listed them as “the structure of the instruments, the nature of the test situation, the orientation to the process and the interpretation of results”. LPAD differs from other learning potential approaches in that it aims to modify each individual, compared to other tests that aim to influence the end product of testing by changing the procedure (Hamers & Resing, 1993). The LPAD is not reaching for Vygotsky’s ZPD, but it focuses on altering the cognitive structures of deprived individuals (Hamers & Resing, 1993; Jensen & Feuerstein, 1987).

The German psychologist, Guthke, and his colleagues developed several types of learning potential tests (Hamers & Resing, 1993). Guthke and Stein (1996) indicated that, because dynamic tests are less strongly related to socio-economic and cultural matters, these tests measure intellectual abilities more fairly than static tests.

Their procedures are closer to the psychometric tradition than to techniques used in other dynamic approaches to testing. They also distinguish between short-term and long-term tests (Grigorenko & Sternberg, 1998). They used repetitions, prompts and systematic feedback during the training between the pretest and post-test. The aim of the test was to increase predictive validity. The training is standardised and restricted to specific hints (Grigorenko & Sternberg, 1998; Guthke & Stein, 1996; Hamers & Resing, 1993). One of their approaches is the testing-the-limits approach, which viewed test performance as a result of dynamic interaction between the individual, the test materials and the test situation. In this approach, manipulation of the test situation can lead to significant improvement of test performance. This approach uses multiple conditions and a variety of aids, varying in verbalisation and feedback. Carlson and Wiedl also contributed to and refined this approach (Grigorenko & Sternberg, 1998; Gupta & Coxhead, 1988; Murphy, 2002).

Campione and Brown, other dynamic assessment theorists, have been influenced by Vygotsky’s ZPD theory of cognitive development. Their tests focus more on the minimum help, in the form of hints or prompts, needed to reach a specific amount of

learning, rather than concentrate on the maximum increase in performance reached through intervention. They also stress transfer as an aspect of learning potential (Hamers & Resing, 1993). The tests of Campione and Brown, like Budoff's test, are quantifiable measures of learning ability as they measure the changes in performance (Minick, 1987).

2.9.3.4 Dynamic assessment in South Africa

Initially, internationally designed dynamic tests were used in South Africa and later other dynamic tests were developed locally.

The approach of Budoff, of administering a standardised traditional measure of cognitive ability in a dynamic test-train-retest manner, was applied in research conducted by Shochet (1986) and Zollezi (1995) regarding student selection. Zollezi's (1995) research indicated that the application of traditional tests in a dynamic manner, is a better predictor of academic performance than static intelligence measures, in the case of disadvantaged students whose ability levels are highly modifiable.

A new dynamic instrument, developed in South Africa, namely the LPCAT (Learning Potential Computerised Adaptive Test), is similar to Budoff's test and has a similar view of dynamic assessment, but uses a test specifically designed for dynamic assessment. Both approaches also utilise non-verbal tasks to determine reasoning abilities (De Beer, 2000c; Murphy, 2002).

Feuerstein's approach was also used in a few research projects in South Africa. The projects of Shochet (1986) and Van Niekerk (1991) will be discussed briefly. Shochet (1986) used traditional and Feuersteinian approaches to assessment in a predictive validity study on student selection in South Africa. Shochet (1986) aimed to determine the effect of enriched conditions on disadvantaged students' performance. After the students wrote a test in the traditional form, training and learning experiences were provided. Thereafter, they were tested again to measure their learning potential. The

results of the research reflected that if disadvantaged students had high manifest levels of functioning, the students were likely to perform at acceptable levels at tertiary institutions, even if they reflected a low modifiability or potential. On the other hand, educationally disadvantaged students, who have a low manifest level of functioning with an ability level that did not appear to be modifiable, did not perform adequately to pass at tertiary institutions. There was evidence in this research that training-enriched testing can help in the selection of disadvantaged students. Van Niekerk (1991) investigated the effects of Feuerstein's instrumental enrichment programme on disadvantaged senior secondary students in South Africa. Her study could not provide evidence that the training or mediation improved verbal or non-verbal reasoning, perceptual speed, mathematical applications, vocabulary or study habits.

De Beer (2000c) pointed out that there was no other reference where the Campione and Brown approach was used in South Africa.

A few dynamic assessment tests were also developed in South Africa; one of them, the Ability Processing of Information and Learning Battery (APIL-B), and another, the Transfer, Automisation, Memory and Understanding Learning Potential Battery (TRAM-1) that was developed by Taylor (Murphy, 2002). De Beer (2000a, 2000b, 2000c) developed the Learning Potential Computerised Adaptive Test (LPCAT), which is also a new dynamic measuring instrument in South Africa. The LPCAT is based on similar principles as the approach of Budoff. Both these tests resort under the measurement or psychometric approach of dynamic assessment. The LPCAT differs from the approach of Budoff, in so far that it uses a test specifically designed for dynamic assessment, compared to Campione and Brown's test who used traditional measures applied in a dynamic manner. The LPCAT will be discussed in more detail in section 4.4.3, as it is one of the tests included in this research project.

2.9.4 Concerns and limitations relating to dynamic assessment

Grigorenko and Sternberg (1998) pointed out that dynamic tests lack a sound psychometric foundation. The biggest concerns are the validity and reliability of learning potential tests. Murphy (2002, p. 35) cited Jitendra and Kameenui (1993) regarding the concerns of dynamic approaches and listed the following issues:

- Construct fuzziness (implying issues of validity)
- Procedural spuriousness (implying issues of reliability)
- Instructional aloofness (implying issues of reliability)
- Instrument inadequacy (implying issues of validity)
- Labour intensiveness (an all-pervasive negative factor in dynamic assessment research)

There are two important psychometric concerns in learning potential research, according to Sijtsma (1993). The first issue is the measurement of change between the pretest and post-test (Grigorenko & Sternberg, 1998; Sijtsma, 1993). Because two different scores are obtained, the reliability of the measure becomes questionable (Boeyens, 1989b). The amount of change is the problematic issue. The change is reflected in the difference in scores of the pretest and the post-test. The pretest score represents the initial state, while the post-test score is obtained after the training intervention. Murphy (2002) indicated that the difference between the pretest and post-test score cannot simply be interpreted as potential.

The second issue is in the ability that was taught, and the possible shift in meaning of the ability due to training. This can possibly influence the construct validity in learning potential testing. Grigorenko and Sternberg (1998), as well as Sijtsma (1993), indicate that the ability to solve certain problems may change qualitatively and/or quantitatively, resulting in different abilities being measured. Hamers and Sijtsma (1993) pointed out that construct validity of dynamic assessment requires more investigation in future.

Guthke (1993) expresses his concerns regarding the theoretical foundations of learning potential tests and points out that learning potential tests “allow only the recording of learning gains but not an analysis of the learning process as such” (Guthke, 1993, p. 55). Guthke also criticised the regular comparison between learning potential tests and conventional intelligence tests, because learning potential tests, in his view, measure more complex problem-solving characteristics than conventional intelligence tests. Guthke (1993), as well as Hamers and Resing (1993), pointed out that more longitudinal studies are required on construct validity of dynamic assessment measures.

Owen (1998) also expressed his concern regarding the psychometric soundness of dynamic assessment measures and referred to it as a psychometrician’s nightmare. He further summarised his other concerns regarding the limitations of dynamic assessment approach as:

- Limited empirical research findings.
- Insufficient work done on technical adequacy .

A final concern regarding dynamic assessment that should be noted is the time that it takes to administer a dynamic test and the resulting cost implications (Boeyens, 1989a). Guthke and Stein (1996) indicated that one of the reasons why dynamic instruments are not used, is that practitioners find these tests awkward and time consuming.

2.10 CHAPTER SUMMARY

In this chapter emphasis was placed on the concepts of intelligence and learning potential as well its the measurement. Reference was also made to the nature nurture debate and the influence of genetic factors and the environment on intelligence.

Psychometric testing and its requirements were reviewed. Thereafter cross-cultural assessment as well as the future of psychometric testing in South Africa was discussed.

The literature review indicated that the utilisation of traditional cognitive measuring instruments is no longer suitable and acceptable to use in the South African situation. The main reason why traditional tests are regarded as unsuitable is that they are static measures that mainly assess crystallised abilities. As all South Africans do not come from the same background, it is necessary to rather consider dynamic instruments that determine the learning potential of a person (Boeyens, 1989a; Foxcroft, 1997; Huysamen, 1996b; Jensen, 1980; Owen, 1998; Shochet, 1986; Van Aswegen, 1997).

In chapter three student selection in South Africa will be discussed and in chapter four the researcher reports on an investigation regarding the use of certain psychometric instruments and other criteria as selection elements that can be included in a selection battery with the aim to predict academic performance of students.

CHAPTER 3

STUDENT SELECTION

3.1 INTRODUCTION

Student selection poses a particular challenge in South Africa, due to the influx of students from educationally disadvantaged backgrounds. As a result of limited state funds there is an urgency to accurately predict students' potential to complete their qualifications.

As the aim of this study is to determine the predictive validity of a selection battery for Technikon students, this chapter is devoted to student selection and the predictors that could be included in a selection battery. Various authors, both nationally and internationally, refer to the importance of the selection of students for higher education (Boeyens, 1989a; Miller, 1992; Rutherford & Watson, 1990; Shochet, 1994; Wilson, 1983; Zaaïman, 1998). The importance of fairness during the selection process will also be discussed. The study seeks to investigate which variables have the highest positive correlation with academic success. The results will contribute towards the identification of a valid selection battery. The various predictors that could have an impact on determining which students have the potential to obtain a diploma will be discussed.

The concept of selection in higher education institutions will be defined in the beginning of this chapter. The components and factors that should be kept in mind when a selection mechanism is designed will also be covered.

Due to the fact that South Africa has a unique multicultural society, with people coming from diverse schooling and socio-economic backgrounds, little direction can be obtained from international practices of student selection. Therefore, an effort was made to focus on South African literature and local research reports in this regard.

3.2 THE CONCEPT OF STUDENT SELECTION

The rationale behind selection of students is the increase in applications to higher education institutions and the diminishing number of places available, resulting from limited funds to accommodate these students (Gardner, 1989; Jackson & Young, 1988; Maxwell, 1996; Nunns & Ortlepp, 1994; Rutherford & Watson, 1990; Uys, 1994; Zaiman, 1998). The South African National Department of Education (DOE) defines higher education in the following way:

Higher Education comprises of all learning programmes leading to qualifications higher than the proposed Further Education and Training certificate or the current Standard 10 Certificate (DOE, 1997, p. 17).

There is legislative pressure on institutions of higher education to find solutions for student selection dilemmas. Section 37 of the Higher Education Act, No. 101 (1997) places the responsibility and accountability of selection on the institution. The Bill of Rights, in the Constitution of the Republic of South Africa, No. 108 (1996), forces institutions to evaluate their selection batteries, by prohibiting unfair discrimination, either directly or indirectly, based on sex, ethnic or social origin, race and culture, amongst others. Finally, the South African White Paper on Transformation of Higher Education (DOE, 1997) also adds further pressure, as it views institutions of higher education as vehicles to achieve equity and address imbalances of the past.

3.2.1 Definitions of student selection

Although various definitions of student selection exist, they emphasise the same aspects. All definitions quoted below were collected from South African literature, to ensure that they are applicable to this project and student selection in South Africa.

- Kriel (2001, p. 102) explained student selection as “the process of choosing the potentially successful student” [italics and underlining added]
- According to Zaaiman et al. (1998, p. 97) “selection for higher education programmes is usually to identify students who will succeed in a specific academic programme”. [italics and underlining added]
- Louw (1993, 1996) defined selection as a process to identify applicants with the necessary potential to complete courses successfully. [italics and underlining added]
- Uys (1994) noted that selection takes place, because more applications than what can be accommodated are received. According to her, the selection of students who will complete the programme is the key factor of selection. [italics and underlining added]
- According to Zaaiman (1998, p. 31), “The main aim of selection to higher education is to identify students who will succeed in a specific academic programme”. [italics and underlining added]
- Jonker (1994) stated that, due to the high costs of tertiary education it is essential that only students with a reasonable chance to complete their studies should be admitted to tertiary institutions. [italics and underlining added]

Italics and underlining were added in above definitions to highlight that all the above authors include the concepts of *completion* and/or *success* in their definitions of selection. This indicates that the selection of students is used with the aim to correctly predict future behaviour in terms of academic success on tertiary level.

3.2.2 Objectives of student selection

According to Zaaiman (1998), the purpose of selection, as stipulated in the South African White Paper on Transformation of Higher Education (DOE, 1997), is to *achieve equity* through developing intellectual capacities to enhance opportunities in life among South African citizens. Fairness and justice are often used together with the term equity. It is thus implied that there should be fair opportunities to enter and succeed in higher education programmes (Mji, 2002).

Through selection, higher education should further contribute to *address the needs of the labour market* in providing high-level competencies to ensure growth in the economy.

Another major objective, as reflected in the definitions provided in section 3.2, is the identification of individuals who have the potential to be academically successful and an indication of those who can complete the qualification (Zaaiman, 1998; Uys, 1994).

3.3 CHALLENGES AND CONCERNS REGARDING STUDENT SELECTION

Access to higher education is regarded as the primary inequality in higher education according to the White Paper on Transformation of Higher Education (DOE, 1997), with specific reference to racial inequalities. Zaaiman (1998) notes that the term “access” refers to the possibility of gaining entry for members of different subgroups.

The main issues and problems in student selection, as summarised by Zaaiman (1998) include, amongst others, the following;

- Increasing numbers of applications.
- Heterogeneous educational backgrounds of applicants.
- Problems in identifying students with potential.

- High failure rates and low retention rates.
- Selection processes lacking transparency.
- Lack of validated selection instruments and policies.
- Too few published research results to assist in the development of new admission practices.

A study done by the Universities Advisory Council reflects the increase in applications that can be expected (Gardner, 1989). This study indicates that the need for places in tertiary education in Southern Africa poses a challenge to student selection.

TABLE 3.1 A COMPARATIVE ESTIMATION OF THE NUMBER OF INDIVIDUALS WHO WILL PASS WITH MATRICULATION EXEMPTION BETWEEN 1990 AND 2010 (Gardner, 1989)

	Number of pupils that passed matric with an exemption in 1990	Estimated number of pupils that will pass matric with an exemption in 2010	Estimated percentage changes in number of pupils between 1990 – 2010
Whites	28 481	24 550	- 14%
Coloureds	4988	12 762	61%
Asians	3831	6084	37%
Blacks	23 463	75 886	69%
Total	60 763	119 282	153%

This table highlights that there will be a drastic increase in candidates who will meet the requirements to attend tertiary education, which will place even more pressure on efficient selection processes within institutions of higher education.

Huysamen (1996a), as well as Rutherford and Watson (1990), pointed out that the largest increase is in the sector where applicants come from the former black education system,

called the Department of Education and Training (DET). They are at a disadvantage in competing for admission to tertiary institutions with candidates who have a better educational background. Rutherford and Watson (1990) indicated that such applicants score lower marks in achievement and aptitude tests, which are the predictors used in selection, and that these measures do not predict future success for these candidates. This poses a further challenge to the selection process.

Zaaiman et al. (1998) indicated that another concern of student selection is that applicant groups are currently more heterogeneous, in terms of educational backgrounds and socio-economic status of the families. Huysamen (1996a), as well as Zaaiman et al. (1998) pointed out that differences in educational backgrounds might still continue despite political changes in the country. Zaaiman (1998, p. 11) indicated that this “complicates the effectiveness and fair selection of students in South Africa”.

Students who attended DET schools can be regarded as educationally disadvantaged, because they did not have adequate opportunities to develop their potential (Huysamen, 1996a; Miller, 1992; Shochet, 1994; Zaaiman et al., 1998). As opportunities are scarce, one should also keep in mind that applicants from the former black education system, the DET, are at a disadvantage in competing for admission to tertiary institutions with candidates who have a better educational background.

Zaaiman (1998, p. 23) stated that:

A student can thus be regarded as having been disadvantaged if he/she has had inadequate access to quality educational services, resulting in a lack of opportunity to fully develop her/his academic potential.

Because of the difference in educational and socio-economic backgrounds of applicants, dynamic assessment measures are regarded as more suitable than traditional static tests. The dynamic instruments are more likely to assess the potential of individuals, whereas traditional measures are said to measure crystallised abilities.

According to Shochet (1994), due to the lack of valid selection instruments, it is a challenge to find appropriate criteria for admission of students. The accurate prediction of future academic performance is thus a requirement for effective selection. Boeyens (1989a) also indicated that more appropriate selection procedures should be developed for disadvantaged students, reconsidering assumptions regarding cognitive ability.

The concerns and challenges raised here serve as motivation to investigate student selection, with the aim to have more efficient and effective selection in future. This will result in valid selection batteries with higher predictive validity, which is what this project aims to accomplish.

3.4 FAIRNESS IN SELECTION

Nunns and Ortlepp (1994), as well as Van Eeden et al. (2001), indicated that fair selection is crucial in South Africa, because of unequal education opportunities in the past. To ensure fair selection, accurate predictors of academic performance should be found (Zaaiman, 1998).

Louw (1993, 1996) emphasised that, because institutions focus on optimal performance, a fair selection process should identify the individuals who have the potential to pass and obtain a qualification. Boeyens (1989a), Jonker (1994) and Shochet (1994) indicated that, to ensure fair selection, provision should be made for students who had inadequate school education, but have the potential to obtain a tertiary qualification. According to Maxwell (1996) selection decisions should be based on fair, impartial and unbiased assessment of relevant qualities to obtain equity in selection.

Nunns and Ortlepp (1994) highlighted that if universities do not want to be seen as using unfair, discriminatory procedures, they should address fairness-related issues. Accurate predictors are therefore essential components in a fair student selection process. They

further indicated the necessity that a tertiary institution's "customers" should have the perception that the selection procedure is fair. Tertiary institution's customers refer to students, the parents and sponsors of prospective students as well as the society at large. Fairness is evaluated from a qualitative or social-ethical perspective as well as a quantitative or empirical view. Fairness depends on the way predictors are used in the selection process. A selection mechanism can be psychometrically valid, but unacceptable to the community according to Zaaïman (1998). It is possible to investigate the validity and therefore fairness of a selection mechanism.

Nunns and Ortlepp (1994) refer to two types of fairness: procedural fairness and distributive fairness.

3.4.1 Procedural fairness

Procedural fairness refers to the perceived fairness of decisions taken in the selection process (Zaaïman, 1998).

Zaaïman (1998) indicated that attention should be given to the following aspects during the selection of students to ensure procedural fairness:

- Consistency of application procedures.
- Accuracy of information.
- Prevention of personal bias.
- The level of representation of all concerned parties (for example institution representatives, student representatives, advisory committees and employers or representatives of business) throughout the process.
- Reflection of current moral and ethical principles.

All of the above are essential to ensure procedural fairness. The accuracy of information and prevention of personal bias can be ensured through research.

3.4.2 Distributive fairness

Distributive fairness refers to whether the right students were selected. The “rightness” of students depends on the context in which selection takes place (Zaaiman et al., 1998). The context is related to the principles of equity and redress within South Africa. Louw (1993, 1996) supports the view of Zaaiman et al. (1998), that fairness in selection should address unequal situations in society. This influences the community’s perception of fairness, and whether they will accept the selection mechanism.

The acceptability of selection mechanisms depends on their perceived fairness (Zaaiman, 1998). If the community and students perceive the selection system as unfair, it can result in anger and resentment against the institution (Nunns & Ortlepp, 1994). Nunns and Ortlepp (1994) pointed out that past disturbances at universities are an indication of the anger and hostility generated by university procedures, rules and regulations which were perceived as unfair.

3.5 DESIGNING A SELECTION MECHANISM

A selection mechanism is regarded as effective if it achieves the desired outcomes. Academic success is one of the main objectives of selection as indicated in the definitions provided. As pointed out earlier, other objectives for higher education that should be kept in mind in designing a selection mechanism are to ensure equity and redress, as well as to address the needs of the labour market (Zaaiman, 1998).

Kotze (1994) indicated that there are four main considerations in formulating selection objectives for a selection mechanism:

- Socio-political considerations refer to the extent to which economic motives and social equality need to be accommodated in selection. This will have an impact on the selection policy. Equality, in the context of selection, is viewed as equal opportunities for gaining access to institutions of higher education.
- Educational considerations relate to the objectives in selection to address the difference in educational background of applicants which impact on their selection.
- Psychological considerations imply that only measuring instruments that will be suitable for the prediction of individual differences in future academic performance should be used.
- Methodological considerations apply to the selection procedures and approaches which should focus on valid and reliable selection decisions.

Kotze (1994) highlighted four steps in the methodological approach towards selection:

STEP 1: The formulation of operational objectives for selection and admittance of students.

STEP 2: The formulation of selection criteria: compulsory and recommended criteria.

STEP 3: Selection of measuring instruments to be used to predict future academic performance, in other words, compiling of a selection battery.

STEP 4: A description of the decision-making process regarding the method that will be used to allocate places.

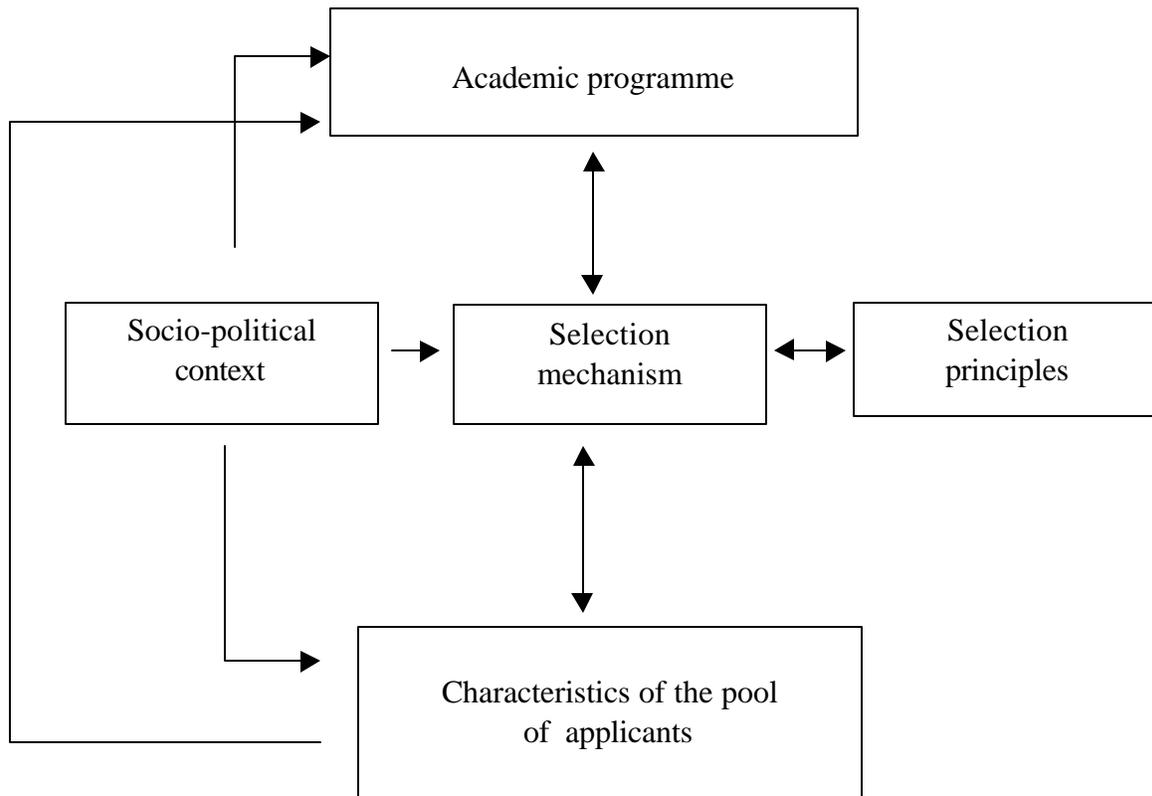


FIGURE 3.1 Factors involved in the development and evaluation of a selection mechanism (Zaaiman et al., 1998)

The selection mechanism, as indicated by Zaaiman et al. (1998) in figure 3.1, indicates the factors that influence the development and evaluation of a selection mechanism. Maxwell (1996) agreed with Zaaiman that the character of selection systems should be based on social and political contexts and further suggested that economic, institutional and philosophical contexts should also be considered.

According to Zaaiman et al. (1998), selection mechanisms include selection tests, selection policy decisions and selection procedures, which will now be briefly discussed.

Selection tests refer to the psychometric tests that are used in the selection process. Shochet (1986) indicated that dynamic rather than traditional psychometric tests should be used, as traditional tests are seen as measures of crystallised abilities resulting from school education, while dynamic measures assess learning potential. Jackson and Young (1988, p. 170) stated that “Effective selection and prediction procedures for tertiary institutions become more crucial as the need to capitalise on the potential of individuals becomes more pressing”. This builds on the argument of Huysamen (1996a) and indicates that potential, rather than crystallised abilities, should be considered, as all applicants do not come from the same background. As mentioned in chapter two, it is not good to use psychometric tests alone. A selection battery should also include other measures, such as matric results for selecting candidates (Huysamen, 1996b; Foxcroft, 1997; Louw, 1993; Van Aswegen, 1997). This project will compare a traditional static test and/or matric results with a dynamic instrument to determine which one has the highest correlation and predictive power to predict academic performance.

Selection policy decisions are used to satisfy selection requirements that cannot be satisfied by selection tests. To be effective and fair, a selection mechanism should be designed for specific courses (Zaaiman, 1998). The demands of the course, like specific pre-requisite matric subjects (for example mathematics and physical science as pre-requisite subjects for engineering courses) should be considered when a selection mechanism is designed (Maxwell, 1996).

Selection procedure refers to the process of selecting students (Zaaiman, 1998). The importance of this was already highlighted in section 3.4.1 under procedural fairness.

According to the Department of Education (1997), an effective selection mechanism will select a high percentage of successful students and reject only a few applicants who could have been successful if they were selected. Jackson and Young (1988), as well as Louw (1993), further support this view by pointing out that effective selection can improve the prediction of success in academic performance.

Ethics also play a role in the selection mechanism, as it involves decisions about the future of individuals, which impact on the community. The efficiency of a selection programme depends on the ability of the selection methods to predict future academic performance in an affordable and sustainable manner (Nunns & Ortlepp, 1994).

3.6 PREDICTORS OF ACADEMIC SUCCESS USED IN STUDENT SELECTION

An approach of considering and investigating the relationship between predictors and subsequent performance is usually used in student selection. Research should identify accurate predictors of academic success (Nunns & Ortlepp, 1994).

The prediction of future behaviour, in the form of tertiary academic performance, is an important element of selection, as the drop out rate at tertiary level is a critical issue (Kotze, 1994; Louw, 1993; Samkin, 1996; Stoker et al., 1985). The extent to which a factor contributes to tertiary academic performance influences the predictive power of the selection instrument. A valid predictor increases the probability of a successful prediction (Zaaiman, 1998). The predictors that are used also contribute to the perceived fairness of the selection battery.

Zaaiman (1998) indicates that, internationally, selection for higher education is mainly based on academic criteria, but indicated that unequal schooling of applicants in South Africa makes the use of school results in South Africa unreliable. Various authors like Boeyens (1989a), Claassen (1990a), Owen (1998) and Retief (1988) indicated that neither matric results nor traditional psychometric tests could be used as fair selection criteria for students from different backgrounds in South Africa. There is thus an urgent need to find alternative predictors that would not unfairly exclude those students from disadvantaged backgrounds (Shochet, 1986). Ayaya (1996) further pointed out that there is no consensus regarding which factors have the most significant influence on future academic performance.

To design a selection battery that is more appropriate as well as valid, the fundamentals regarding underlying principles of ability, potential and mental development is required (Boeyens, 1989a). According to Boeyens (1989a), ability, as measured with traditional psychometric tests, is relatively independent of learning potential. Louw (1996), Nunns and Ortlepp (1994), as well as Shochet (1992), supported the view of Boeyens that potential ability should rather be assessed. As indicated earlier in this report there is a growing awareness that psychometric test results gain meaning if they are integrated with other information (Ayaya, 1996; Foxcroft, 1997; Nunns & Ortlepp, 1994; Skuy, Zolezzi, Mentis, Fridjon & Cockcroft, 1996). Huysamen (1996b) supports the view of Foxcroft (1997) that if demographic data and psychological test results are used together, it could result in a fair and unbiased admission procedure at tertiary institutions. Foxcroft (1997) also notes that there is evidence that South African test developers and users shifted their approach from unitary testing to multi-method assessment.

Ayaya (1996), Brown (1994), Fourie (1991), Jackson and Young (1988), Jonker (1994), Kotze (1994), Nunns and Ortlepp (1994), as well as Zaaïman (1998), pointed out that intelligence is not the only factor that could influence academic performance and indicated that certain non-cognitive factors also play a role.

The following are examples of factors that can play a role in future academic success:

- Gender
- Age
- Language proficiency
- Motivation
- Locus of control
- Satisfaction with the institution
- Ability to make effective decisions
- Health
- Emotional balance

- Personality variables
- A need for achievement
- Study habits and attitudes
- Quality of the tertiary academic environment

Russel, Persing, Dunn and Rankin (1990) suggest that it is necessary to understand the role that each variable plays in the selection process, as all predictors are not equal in terms of their predictive value. They further point out that unequal predictors should then not be given equal weight in the selection decision.

It is clear that quite a number of possible predictors can be used in selection. In this chapter the researcher does not aim to provide a detailed discussion of all the factors that can possibly influence academic performance. Neither does she discount nor discredit the influence of the factors mentioned, but will suffice to mention that other factors can also influence academic performance. The aim is rather to specifically assess the predictive validity of matric results and a dynamic psychometric test compared with a traditional static measure of cognitive ability. As the project also aims to investigate correlations between academic performance and certain biographical data, as well as the bridging course status of students, a discussion of these predictors will follow now.

Van Zyl, Terblanche and Jacobs (1992) stated that if valid predictive selection measures could be found, it could prevent personal failure for students and save money for institutions. Jackson and Young (1988) pointed out that the prediction of success in higher education continues to be a pressing problem and indicated that an effective selection system might improve the prediction of success.

These statements necessitate investigations, like the present research project, to find selection models that will be more suitable to base selection decisions upon.

3.6.1 Matric results

Matric results represent a sample of academic behaviour and were traditionally used, together with psychometric tests, to predict academic performance at tertiary institutions (Huysamen, 1999; Huysamen & Raubenheimer, 1999; Louw, 1996; Rutherford & Watson, 1990; Zaaiman, 1998; Zaaiman et al., 1998). This is no longer suitable, as applicants come from heterogeneous educational backgrounds. Shochet (1986) agrees that the use of matric results for admission of applicants to tertiary institutions is suspect, because of evidence of disadvantages in black education. Louw (1996) stated that the education system of the Department of Education and Training (DET) provides an inferior schooling for some applicants. The matric results from the former Department of Education and Training (DET) were regarded as an unreliable predictor of future academic performance of disadvantaged students (Rutherford & Watson, 1990; Zaaiman, 1998).

3.6.1.1 Shortcomings of matric results as predictors of academic performance

According to Boeyens (1989a), the problems with matric results, in terms of their suitability for the prediction of academic success, are as follows:

- In industrialised societies most students go through an education that optimises their performance. In South Africa this is not the case for all candidates.
- Research showed that academic achievements in “white” schools are better predictors for academic success at tertiary level than achievements at “black” schools.
- In South Africa the quality of school education varies.

Huysamen (2000, p. 146) indicated the following with regard to budget allocations in “black” and “white” schools, “In 1988 – 1989, R656 was spent per black child on education as opposed to the R2 882 per white child”. Differences are mainly the result of

teachers' salaries at "black" and "white" schools, based on their qualifications. Huysamen (2000), citing Christie (1991), pointed out that 32 percent of black teachers were not even matriculated, as opposed to white teachers who have all matriculated with 32 percent holding university degrees.

It will be inaccurate to base selection decisions on matric performance, as it is clear that everybody does not have the same quality schooling (Huysamen, 2000).

3.6.1.2 Research findings regarding the suitability of matric results as predictors of academic performance

Various views exist on the suitability of matric results as predictors of academic success. Some researchers single out matric results as the most reliable and useful indicator of academic success, while others proved that matric results do not correlate with future academic performance.

Fourie (1991) in a study she conducted in 1987, reported that students who passed all subjects at the end of their first year had significantly higher or better matric symbols. The discriminant analysis that was used reflected that 66.8 percent of the successful students performance could be ascribed to their matric results. This study was, however, conducted on a group of Afrikaans speaking students at RAU, and it is likely that they probably attended schools that prepared them appropriately for tertiary education. In another study conducted on nursing students at various campuses of Natal College (Uys, 1994), it was reported that the influence of school learning seems to only have an impact on first-year academic achievement. Research conducted by Samkin (1996) on 511 full-time accounting students (matriculants from DET and HOD [House of Delegates]) at the University of Durban-Westville, indicated a linear relationship ($r = 0.53$ and $p < 0.05$) between students' overall matric results and their performance in Accountancy I. In a study including 6527 students, which Stoker et al. (1985) conducted, the aggregate score in the matric examination proved to be the strongest predictor of success at university. A

study conducted by Van Eeden et al. (2001) showed that school performance in Mathematics ($r = 0.25$; $p < 0.01$), Science ($r = 0.28$; $p < 0.01$) and English ($r = 0.33$; $p < 0.01$) were the best predictors of average first year performance in science courses. Huysamen and Raubenheimer (1999, p. 176) also reported that “the almost routine rejection of matriculation marks as a wholly inadequate predictor ... of DET ex-students is rather premature”.

Other researchers showed the limited predictive value of matric results. Shochet (1986) reported that, while there was a statistically significant correlation between matric results and university success for educationally advantaged students (matric results explained 30 percent of the variance of students' average results), the same study reflected that DET matric results did not predict the academic success of disadvantaged students. Miller (1992), Skuy et al. (1996), as well as Zaaiman (1998), also came to the conclusion that South African matric results are not valid or reliable, as a result of differences in scholastic education of candidates.

3.6.2 Traditional psychometric tests

The rationale behind using intelligence tests for selection is that the tests measure cognitive abilities and because, it is assumed that there is a relationship between intelligence and academic achievement (Shochet, 1986), these tests are used to predict academic success. All tests should meet certain requirements in terms of reliability, validity and standardisation which was already discussed in section 2.5.1. The extent to which the instruments used in this study meet these criteria will also be reported on.

As indicated in chapter two, traditional psychometric measures only represent the current abilities and fail to give an indication of the potential that individuals have to succeed. Current abilities are seen as the result of opportunities to learn in home and school environments. Shochet (1986) further indicated that, because traditional tests have a

static view of intelligence and because they do not take the modifiability of cognitive abilities into account, they fail to distinguish between current and potential performance.

3.6.2.1 Shortcomings of traditional psychometric tests as predictors of academic performance

Internationally, Budoff (1987b) and Wilson (1980) indicated that traditional IQ tests often misclassify children. Budoff (1987b) argued that IQ tests are unable to distinguish between children with a low IQ score who are unable to benefit from experience, and those who are able to benefit from experience. According to Babad and Budoff (1974), it is possible to find low IQ children who are competent problem solvers in their non-school environments. These children have therefore learned from their environments. If students are provided with a learning environment at tertiary level, they might succeed academically.

Locally, Boeyens (1989a) indicated that the scores achieved in psychometric tests could not be interpreted as if they represent the ceiling of a person's ability, because all candidates did not have the same opportunities to develop their abilities. Due to unequal educational opportunities in South Africa, some assumptions underlying psychometric testing for industrial and educational selection are questionable, according to Boeyens (1989a).

3.6.2.2 Research findings regarding the suitability of traditional psychometric tests as predictors of academic performance

In one study two sub-tests of the AAT, English Reading Comprehension and Numerical Comprehension, were administered to a group of 511 first-year accounting students at the University of Durban-Westville. The scores obtained in these sub-tests did not prove to be accurate predictors for Accounting I results in 1992 for DET students (Samkin, 1996).

Nunns and Ortlepp (1994) used the Mental Alertness Test (MAT) in exploring predictors of academic success in Psychology 1 and found that the MAT correlated statistically significant ($r = 0.37$; $p < 0.0001$) with Psychology 1 for educationally advantaged students, but not for educationally disadvantaged students. Research done by Skuy et al. (1996) revealed that the MAT, a static measure of intelligence, was a statistically significant predictor for the educationally advantaged group ($r = 0.75$; $p < 0.01$), but did not predict academic success for educationally disadvantaged students. Van Eeden et al. (2001) correlated GSAT and SAT with the average first year performance of science and technology students coming from a disadvantaged educational background. Their results reflected that the GSAT full scale correlated ($r = 0.24$) statistically significant at $p = 0.01$ with average first year results. The correlation coefficients between SAT calculation ($r = 0.21$) and SAT Spatial 3D ($r = 0.18$) were also significant at $p = 0.05$ level.

3.6.3 Learning potential assessment

As indicated earlier, traditional psychometric tests are generally regarded as unsuitable for predicting academic performance fairly in the multicultural context and diverse background of individuals in South Africa.

Guthke and Stein (1996) indicated that learning potential test results provide more information than static measures of intelligence. Guthke and Stein (1996) further reasoned that static tests are suitable if one is only interested in the current intellectual status of a person, but if predictions must be made for individuals from different backgrounds, dynamic instruments are more suitable. The modifiability of the individual's cognitive ability indicates the learning ability or intelligence of the person (Nunns & Ortlepp, 1994) and has important implications for the prediction of academic performance for multicultural groups.

There is optimism regarding the use of learning potential assessment in student selection in South Africa, because it is viewed as more fair to use in situations of unequal

educational opportunities (Boeyens, 1989a; Nunns & Ortlepp, 1994; Skuy et al., 1996; Zaaïman, 1998). Shochet (1986) also pointed out that the measurement of learning potential, which assesses the modifiability of an individual, might be a “better” and “fairer” predictor of success at a tertiary qualification level than static measures of intelligence, and therefore a more valid predictor of academic success in multicultural selection.

Allen, Woodward, Jones, Terranova and Morgan (1990, p. 279) indicated that:

There is a conceptual and empirical rationale for new research aimed at determining the value of learning rate measures for the prediction of success in training or educational courses.

They further state that:

research should (a) provide a comparison of the predictive validities of traditional psychometric measures and measures derived from samples of learning under controlled circumstances ... (b) determine the extent to which direct measures of learning abilities complement psychometric measures in the prediction of successful learning.

3.6.3.1 Shortcomings of learning potential tests as predictors of academic performance

Boeyens (1989a) and Zaaïman (1998) indicated that it is not practical to use dynamic tests in mass application for student selection purposes, because it is time-consuming and resource intensive.

Other concerns and limitations regarding dynamic assessment measures were already raised in section 2.9.4. Various researchers indicated that the use of dynamic assessment in student selection requires more investigation (Murphy, 2002; Shochet, 1986; Zolezzi, 1992).

The biggest concern regarding learning potential tests is the psychometric soundness of these measures. With regard to the reliability, the two tests scores that are obtained raise questions regarding the reliability of learning potential measures (Boeyens, 1989a; Grigorenko & Sternberg, 1998). There are also concerns regarding the construct validity of psychometric tests (Grigorenko & Sternberg, 1998; Murphy, 2002).

Further concerns are expressed regarding the time and cost implications in administering a learning potential test (Boeyens, 1989a).

3.6.3.2 Research findings regarding the suitability of learning potential tests as predictors of academic performance

Kotze, Van der Merwe and Nel (1996) used the APIL (Ability, Processing of Information and Learning Battery) on 2336 students at RAU (Randse Afrikaanse University). Their study revealed a statistically significant correlation between total APIL scores and academic performance of students. Research conducted by Nunns and Ortlepp (1994) also revealed positive and statistically significant correlations between the Conceptual Reasoning Test (a measure of learning potential) and academic performance of educationally disadvantaged students ($r = 0.48$; $p < 0.01$) and for educationally advantaged students ($r = 0.44$; $p < 0.0001$). They did, however, warn that the sample size of the educationally disadvantaged students was small and that these findings should be used with caution and that more research is required in this regard. A study conducted by Skuy et al. (1996) found that the Learning Process Measure (LSP), a learning potential test, correlated ($r = 0.56$; $p < 0.5$) statistically significantly with the average June score of students registered for a bridging course at the University of Witwatersrand. These findings therefore indicate that learning potential tests can be regarded as suitable predictors of academic success.

Boeyens (1989b) reported on research, which he conducted at a PROMAT (Project Matric) college. PROMAT aims to provide an enriched educational environment to previously disadvantaged students.

Shochet (1986) implemented the Feuerstein concept of mediation and testing through a test-coach-test principle. His findings were based on a sample of black students from a disadvantaged background. His research showed that disadvantaged students whose manifest level of functioning was high to begin with, are likely to succeed even if they are not very modifiable. Disadvantaged students, with a lower manifest level of functioning and who do not appear to be modifiable, are more likely to fail.

In contradiction to this finding, Van Eeden et al. (2001) reported on a study that they conducted on 224 students from science and technology departments at ML Sultan Technikon in Durban. The LPCAT was included in the battery of tests that they used. Average first-year results over all subjects were used as criterion. In this study both the pretest ($r = 0.18$) and post-test ($r = 0.13$) scores of the LPCAT did not reflect a statistically significant correlation with the criterion. They did, however, conclude that the aim of the LPCAT is to measure modifiability and should be used to identify students who will benefit from support programmes. In another research project conducted in 2003, where the LPCAT was used, there was a statistically significant correlation ($r = 0.23$; $p < 0.05$) between first-year average of technikon students and LPCAT post-test (De Beer, 2003).

3.6.4 First-year results

Various authors on student selection place emphasis on the completion of qualifications at institutions of higher education (Louw, 1993, 1996; Nunns & Ortlepp, 1994; Uys, 1994; Zaaiman, 1998; Zaaiman et al., 1998). Ayaya (1996) and Huysamen (2000) also investigated the influence of first-year results on successive years.

3.6.4.1 Shortcomings of first-year results as predictors of academic performance

Wilson (1983) argued that if a single subject or only the first-year's academic performance is considered, it is a short-term criterion for success. He suggested that criteria that reflect a student's long-term performance should rather be considered.

The emphasis on the completion of a course indicates that first-year results are viewed as inadequate to use as a criterion. Academic performance at first-year level cannot automatically be regarded as proof of successful completion of a qualification (Huysamen, 2000). Louw (1996) also recommended that future research regarding the success of students should rather use the full course as a criterion.

Huysamen (1999, 2000) pointed out that poor predictability of first-year results could be a result of inadequate command of the language, and that students cope better with increased language ability later in their studies.

3.6.4.2 Research findings regarding the suitability of first-year results as predictors of academic performance

In a study that Huysamen conducted in 1996 on 1133 students at the University of Orange Free-State, the following correlation scores were obtained. MSPT correlated at $r = 0.21$ for black students and at $r = 0.64$ for white students with MPM (Mean percentage mark, numerical value indicate the relevant year) at a statistically significant level. The correlations between MPM 1 and MPM 2 for black students were $r = 0.557$ and $r = 0.779$ for white students. First-year performance correlated higher with second year results, than matric results with MPM 1, for both black and white students. The correlation between MPM 3 and MPM 2 for black students was $r = 0.638$ and for white students $r = 0.775$. There were even indications of a higher correlation between the second-year and third-year results for black students while the correlation between

second-year and third-year results of white students decreased. He pointed out that these findings were in agreement with findings of Wilson in the United States of America (Huysamen, 2000).

3.6.5 Biographical factors

There are various references in the literature regarding the influence of biographical factors, such as home language or language proficiency and gender on academic performance in higher education (Huysamen, 1999, 2000; Huysamen & Raubenheimer, 1999; Van Rooyen, 2001; Van Eeden et al., 2001). The effect of language proficiency and gender will be investigated in this research project.

3.6.5.1 Shortcomings of biographical factors as predictors of academic performance

Kriel (2001) indicated that it would be unconstitutional to use biographical factors, such as age and gender, as part of a selection battery to deny or grant applicants access to tertiary institutions in South Africa.

If a student completes predictor variables, such as matric examinations and cognitive tests, or criterion variables, like academic performance, in another language than their mother tongue, this can affect the reliability and validity of the scores obtained in these variables (Huysamen, 1999). Biographical factors are therefore not regarded as important predictors, but they will still be monitored in this study, as it remains important to find combinations with the best predictive value for student selection.

3.6.5.2 Research findings regarding the suitability of biographical factors as predictors of academic performance

With regard to gender, it is interesting to note that research done by Everette on 1695 full-time students in Australia during 1987 proved a consistent tendency for women to have higher pass-rates than men (Everette, 1991). Similar findings were obtained in the research conducted by Ayaya (1996) in South Africa. Zaiman (1998) referred to Brusselmans-Dehairs and Henry (1994) who found that there are indications that boys generally perform better than girls in mathematics. In a research project conducted by Huysamen (2001), men's mean percentage mark (MCPM) correlated higher with the matric symbol point total (MSPT) than that of women. Men's MCPM correlated with MSPT at $r = 0.31$ compared a correlation of $r = 0.18$ between MCPM and MSPT for women.

In a study conducted by Van Rooyen (2001), home language was found to be a significant predictor of academic performance, explaining 10.8 percent ($p < 0.01$) of the variance in MCPM (mean of percentage marks). Agar's (1990) study revealed that students attribute their success or failure in academic performance to language skills and abilities or deficiencies.

3.6.6 Bridging course results

The South African White Paper on Higher Education (DOE, 1997) stressed the necessity of bridging programs within higher education. These programs aim to bridge the gap between school and higher education institutions (Zaiman et al., 1998). It also succeeds in narrowing the gap between students from different educational backgrounds (Huysamen, 1996a). High failure rates and unequal educational backgrounds of applicants necessitate higher education institutions to consider implementation of relevant support programs for unprepared students (Ayaya, 1996; Van Rooyen, 2001; Zaiman et al., 1998).

3.6.6.1 Shortcomings of bridging course results as predictors of academic performance

No specific reports were found on any shortcomings in bridging course results as predictors of academic success in the literature study conducted.

3.6.6.2 Research findings regarding the suitability of bridging course results as predictors of academic performance

Zaaiman (1998) did research on a foundation course at the University of Fort Hare. The course aimed to develop mathematical and physical science skills to prepare students for scientific and engineering courses. Her research findings proved a significant correlation between the foundation year as selection tool and later academic performance. Ayaya (1996) and Huysamen (1996a) also reported that bridging programmes could lead to improved performance of students.

3.7 CHAPTER SUMMARY

In this chapter, the researcher aimed at defining selection in the context of admission to tertiary institutions. Attention was given to the objectives in student selection as well as challenges and concerns in this regard. A discussion on the issue of fairness in selection was also included in the chapter, whereafter a selection mechanism was discussed.

The predictors which can be used in student selection because of their impact on the predictive validity of student selection, were also referred to.

The predictive validity of the PIB and the LPCAT, as possible alternatives to traditional psychometric tests, will be investigated and reported on in the next chapter.

CHAPTER 4

RESEARCH PROCESS

4.1 INTRODUCTION

In this chapter the empirical part of this research project is described. This project is a quantitative study reflecting a longitudinal comparative analysis between various predictors and tertiary academic performance of students. Empirical research data were collected between 2000 and 2003.

This chapter contains a discussion and explanation of the research process, the measuring instruments as well as the criterion measure that was used in the study. The sample used and the procedures of data collection will also be described.

4.2 RESEARCH DESIGN AND STATISTICAL METHODS

A correlational research design and statistical methods was used for this study. Correlation studies are used to investigate relationships between variables and to make predictions, as they reflect the relationship between the independent variables (x) and the dependent (y) variables. Correlation analysis and regression analysis are the two statistical methods that will be used to quantify and describe the possible relationship between the predictors and criteria described in the previous chapter. Both the commonly known correlation coefficients were calculated, namely Pearson's product moment correlation coefficient and Spearman's rank order correlation coefficient.

Pearson's product moment correlation (r) measures the strength of the linear relationship between two variables. Spearman's rank correlation (r_s), also referred to as rho or ρ , the Greek letter for r , is similar to Pearson's moment correlation, but replaces observations

with their ranks. The Spearman correlation coefficient is more meaningful, as its interpretation does not require that a relationship must be linear (Diekhoff, 1996; Tredoux & Durrheim, 2002). Spearman's r_s is more robust than Pearson's r , because it is based on ranks and not actual scores, and is therefore less influenced by outlying values (Cohen, 1969; Tredoux & Durrheim, 2002).

Correlation coefficients are expressed as r and range between $r = -1$ and $r = +1$, reflecting the strength of a relationship between variables (Wegner, 1993). Correlation coefficients can also be interpreted in terms of effect sizes. See table 4.1 in this regard. Both the strength and direction of a correlation value are indicated. The closer the coefficient is to one, the stronger the relationship that exists between the variables, and in such cases values of the independent variable can confidently be used to predict the dependant variable. The closer the score is to zero, the weaker the relationship. The direction of the correlation coefficient is indicated by the negative (-) or positive value (+) that accompanies the r value (Wegner, 1993).

Wegner (1993) pointed out that a low correlation does not necessarily imply the variables are unrelated, but simply that the relationship is not linear. A non-linear relationship might exist. He further points out the statistical principle that a correlation does not imply a cause and effect relationship between variables, but only an observed association between the variables.

Correlation values are interpreted in two ways, namely in terms of statistical significance and practical significance.

The statistical significance of statistical values is very important. Statistical decisions are made on the basis of probability and are always uncertain (Cohen, 1969; Tredoux & Durrheim, 2002). Statistical significance is usually interpreted in terms of the p (probability) value.

The following guidelines, regarding the interpretation of the statistical significance, are used in this project:

- * significant at $p < 0.05$,
- ** significant at $p < 0.01$ and
- *** significant at $p < 0.001$.

To determine the practical significance, the effect sizes can be evaluated. For the differences between means, effect sizes are reported as a standardised index, which expresses the difference (d) between the means in standard deviation units (Cohen, 1969; Tredoux & Durrheim, 2002). Correlation coefficients can be directly interpreted as the effect size. Effect size provides an index for practical significance, rather than statistical significance, reflecting the practical usefulness of the statistical scores obtained. The American Psychological Association (APA) indicated that it is regarded as good scientific practice to report effect sizes and urged researchers to report and interpret effect sizes (Cohen, 1969; Kirk, 1996; Tredoux & Durrheim, 2002).

TABLE 4.1 INDICES FOR INTERPRETATION OF EFFECT SIZES

Statistical technique	Effect size symbol	Interpretation
Comparing differences between means	d	0.2 = small effect 0.5 = medium effect 0.8 = large effect
	$d = \frac{M1 - M2}{s}$	
Correlations	r	0.1 = small effect 0.3 = medium effect 0.5 = large effect

Correlation analysis is an appropriate statistical technique to use in this study, as the study aims to measure the relationship between certain variables as stipulated in the aims of the study in section 1.3. Correlation analysis will be conducted first, to confirm whether the variables are related. Thereafter, regression analysis will be performed to determine the degree to which variables explain the variation in the criteria. Regression

analysis is used to quantify the relationship between one or more independent variable(s) and a dependant variable. “Regression analysis aims to find a linear (i.e. a straight line) function which best fits the actual observations” (Wegner, 1993, p. 302). In this study stepwise multiple regression analysis will be used, as there are various independent variables that will be used to predict values of the dependant variable, namely the academic performance of students. Various models, referring to combinations of independent variables, will be subjected to the multiple regression analysis to determine which variables explain the highest proportion of variation in the criteria within that model

The following table will indicate the criteria and predictor variables of this study. They will be discussed in detail in section 4.4 and 4.5.

TABLE 4.2 PREDICTORS AND CRITERIA MEASURES THAT WERE USED IN THIS RESEARCH PROJECT

Predictors (Independent variables)	Criteria (Dependent variables)
Matric symbol point total (MSPT)	Average of first semester first year
Matric English score	Average of second semester first year
Commercial subjects in matric status	Average first year academic performance (MCPM 1)
LPCAT pretest score	Average of first semester second year
LPCAT post-test score	Average of second semester second year
PIB Reading Comprehension sub-test	Average of second year academic performance (MCPM 2)
PIB Calculation sub-test	Average of first semester third year
PIB Mental Alertness sub-test	Average of second semester third year
Bridging course status	Average of third year academic performance (MCPM 3)
Gender	Number of semester courses obtained over three years
Average of first semester first year	
Average of second semester first year	
Average first year academic performance (MCPM 1)	
Average of first semester second year	
Average of second semester second year	
Average of second year academic performance (MCPM 2)	

4.3 SAMPLE

The population included all students from the National Diploma: Human Resources Management who registered for Personnel Management I for the first time in 2000 at Technikon Pretoria Nelspruit campus. Students who repeated Personnel Management I in 2000 were excluded as they were already admitted to the National Diploma in 1999. A sample of convenience was therefore used and included 72 students (N=72).

Only students of the National Diploma: Human Resources Management were included, because other courses do not consist of the same subjects. Different subjects will influence the mean percentage mark (MCPM) of students and the MCPM of students will not be comparable. If students with different combinations of subjects are included, it can influence correlation and regression results and influence the reliability and validity of the statistical results.

All students in the sample matriculated at DET (Department of Education and Training) schools and are therefore from the previously disadvantaged group. Other detail regarding the gender distribution of the sample will be provided below.

The gender distribution of the sample was as follows:

TABLE 4.3 GENDER COMPOSITION OF THE SAMPLE

Gender	Number of students	Percentage of sample (%)
Males	44	61
Females	28	39
Total population	72	100

To be able to report on the effect of a bridging course on the academic performance of students, the sample was also divided into students who did the bridging course and those who did not (see table 4.4). The bridging course was developed to bridge the gap between matric and tertiary education, with the aim of improving academic performance and increasing the overall pass-rate in all the subjects of the national diplomas.

The bridging course consists of eight subjects, namely:

- Communication: English
- Business Management
- Economics
- Business Law
- Introduction to Marketing
- Introduction to Human Resources Management
- Accountancy Skills
- Computer Skills

The aims of the bridging course are:

- To improve the students' language ability.
- To introduce them to commercial subjects, as most candidates applying for commercial courses have not done any commercial subjects at school level.
- To help them with the adjustment to tertiary education.

Students register for the bridging course because of the following possible reasons:

- They were not selected for the National Diploma.
- They applied when the National Diploma had no vacancies to admit more students.

- They applied after the closing date of the National Diploma and opted to do the bridging course.

Students who pass the bridging course get automatic access to the National Diploma. They do not go through the normal selection procedure, which is based on their matric results and PIB scores, as it is assumed that the bridging course that they completed bridged the gap between school and tertiary education.

TABLE 4.4 PERCENTAGE OF STUDENTS IN THE SAMPLE WHO COMPLETED A BRIDGING COURSE

Classification relating to the bridging course	Number of students	Percentage of sample (%)
Students who did a bridging course	26	36
Students who did not do a bridging course	46	64
Total sample	72	100

4.4 THE INDEPENDENT VARIABLE MEASURES

The following measures, which were used as independent variables in this study, will be discussed in detail:

- Matric results
- Potential Index Battery (PIB) scores
- Learning Potential Computerised Adaptive Test (LPCAT) scores

The other independent variables that will be used in this project (listed in table 4.2) included the individual matric subjects, namely English and commercial subjects. The

bridging course and gender, as independent variable, were already discussed in section 4.3. The tertiary academic performance of students on first and second year level are also listed, as it was a secondary aim of this project to determine the correlation between academic performances in the different years over the three-year period. This will be discussed in section 4.5 and 5.3.3.

This research project's aim is to find the predictor or combination of predictors (independent variables), which is most accurate in predicting future academic performance. It must be noted that investigating the predictive validity of the psychometric tests and other predictors do not constitute unqualified support for using it in student selection. Nevertheless, the predictive validity of any test should be determined before using it for selection purposes, as indicated in section 2.5.1.2. This project rather aims at determining whether the two psychometric tests, matric results and/or other predictors predict future academic success of students most accurately. Each one, or all, of the predictors in table 4.2 will be evaluated in terms of their predictive validity in predicting future academic performance. This will be done by means of multiple regression analysis.

The major independent variables, namely the two psychometric tests and matric results, will now be described. Specific attention will be paid to the validity and reliability of these measures.

4.4.1 Matric results

A number of researchers, nationally and internationally, have used final school marks in their investigations of predictors of future academic performance (Huysamen, 1999; Jonker, 1994; Kriel, 2001; Louw, 1996; Rutherford & Watson, 1990; Shochet, 1994; Stoker et al., 1985; Van Rooyen, 2001; Zaaiman, 1988).

A matric certificate is a minimum prerequisite for registration at tertiary institutions. All applicants have to submit a copy of their matric results together with their applications. The matric results of the students in this sample were obtained from their individual files at the administration department of Technikon Pretoria, Nelspruit satellite campus. The matric symbols of each student were individually recorded and converted to numerical values according to table 4.5.

Forty seven students (66 percent) of the sample took six or more subjects, while the other 25 students (34 percent) took only five matric subjects. Where students had six or more subjects, the subjects with the lowest symbols were disregarded and only the five best symbols were captured. This procedure was clarified with Huysamen (personal communication, 22 July 2003) who used a similar approach. Fourie (1991), as well as Van Zyl et al. (1992), also applied the principle of the subjects with the highest symbols in calculating the MSPT in their research projects on student selection.

The Swedish formula (Huysamen, 1999, 2000, 2001, 2002; Huysamen & Raubenheimer, 1999; Jonker, 1994; Kotze, 1994; Kriel, 2001; Louw, 1993; Shochet, 1994; Van Rooyen, 2001; Zaaiman, 1998) was used in awarding numerical values to matric symbols and to calculate a numerical matric value for each student. This mark is called the matric symbol point total, generally known as the MSPT.

TABLE 4.5 SWEDISH FORMULA FOR QUANTIFICATION OF MATRIC SYMBOLS (Huysamen, 2002; Kotze, 1994; Louw, 1993)

Matric symbol obtained in subject	Higher grade	Standard grade
A	8	6
B	7	5
C	6	4
D	5	3
E	4	2
F	3	1
G	2	0

The literature review reflected that the Swedish formula is often used in similar research projects (Huysamen, 1999, 2000, 2001, 2002; Huysamen & Raubenheimer, 1999; Jonker, 1994; Kotze, 1994; Kriel, 2001; Louw, 1993; Shochet, 1994; Van Rooyen, 2001). The application of the same procedure, of transforming matrix symbols into numerical values, makes this study's results comparable with other research results on predictive validity of student selection.

4.4.2 Potential Index Battery (PIB)

The PIB was included in this study, as it is the psychometric instrument that is currently used for student selection purposes at Technikon Pretoria.

The PIB was developed as a joint project between the National Productivity Institute (Erasmus & Minnaar, 1995) and Potential Index Associates. It was developed for assessing people's suitability for certain positions, in other words to be used for personnel or staff selection. The PIB contains 25 sub-tests, called indexes, measuring both cognitive skills and interpersonal skills. It is aimed at South African respondents of 16 years and older with a minimum of standard six (Grade eight) level of education. Erasmus (1995, p. 2) describes the PIB as "suitable for the establishment of potential in areas of human performance regarded as critical for the prediction of performance and success (in the workplace)". The PIB measures intellectual, social and emotional competencies. The test is computerised and easy to administer (Erasmus & Minnaar, 1995).

It is necessary to note that the term "potential" in the title of this test can be ambiguous. Murphy (2002) pointed out that the use of the word "potential" in a test title could be misleading with regard to dynamic instruments and stressed that practitioners should be informed about the differences in these terms. She indicated that the word "potential" in some test titles aims to convey the same meaning as ability, although these terms are not synonymous.

The PIB is not regarded as a potential test resorting under dynamic tests. The PIB is regarded as a static measure of intelligence, as no training intervention takes place during the testing process and there is also no pretest and post-test. The PIB should thus not be confused with dynamic measures of ability, just because it has the word “potential” in the test title (Murphy, 2002). Potential in PIB terms is defined as a combination of capacity and competence, and takes a multi-dimensional profile of a person, rather than a single index, into account (Erasmus, 1997).

Erasmus (1997) highlighted two interrelated and interdependent principles that were used in the development of the PIB to ensure the fairness of the assessment tool. They are:

- The instrument must be fair to the person.

Erasmus (1997, p. 3) believes that the “fairness of a tool goes beyond the culture of a group of people or a specific society or community as such”. He pointed out that in assessing the potential of a person, the assessment should also disregard the environment that might influence the potential being assessed. He clarified this by saying “if both individuals have the same potential, the instrument must reflect this fact without any noise in the channel of communication” (Erasmus, 1997, p. 3).

- The instrument must be fair to the situation.

This principle refers specifically to the workplace. Erasmus (1997) pointed out that if Einstein was “discovered” based on the first principle of fairness to the person, and he were placed in an environment where English was the communication medium, with the aim to utilise his potential effectively, it would not be fair to Einstein or the situation. The PIB, as personnel selection tool, therefore assesses the situation while also assessing the individual’s suitability for the situation. In student selection the developers suggest that the profile of an ideal candidate should first be determined. The applicants’ abilities should then be assessed to determine their suitability for the course they applied for. The

suitability of a candidate depends on whether he/she has the required abilities to pass the National Diploma.

Over and above the fairness aspects which Erasmus pointed out, other aspects of the PIB that result in the fairness of this test according to Erasmus (1997) are:

- Validation studies conducted on representative groups, from a broad spectrum of South African backgrounds, indicated no significant difference between race, gender, background and potential measured by the PIB. This is regarded as an indication of the culture fairness of the measure.
- As a result of computer scoring, the evaluation process is objective, easy-to-administer and fair, thus resulting in evaluation fairness (Erasmus, 1997; Erasmus & Minnaar, 1995).

4.4.2.1 Administration of the PIB

The PIB can either be used as a pen-and-paper test, or in its computerised version. The testing duration of the PIB depends on the selection of indexes that will be used.

All PIB scores are automatically converted to a five-point scale by means of a COMPUCON computer programme. The following general interpretation applies to the five-point scale (Erasmus & Minnaar, 1995):

5	=	Excellent
4	=	Very good
3	=	Good
2	=	Fair
1	=	Poor

Norms were developed and these apply to all respondents regardless of race, gender and culture (Erasmus & Minnaar, 1995).

4.4.2.2 Standardisation of the PIB

Erasmus (1997, p. 4) criticised the usual standardisation methods of falling short “in real sensitivity of the actual influences of the expanding (cultural) diversity and accelerated change to which man and society are currently exposed”. He further pointed out that the “traditional way of standardisation is often justified as a way in which ‘cultural diversity’ is taken into account, through using separate norms for the different culture groups”.

Erasmus (1997) pointed out that often, once standardisation has been completed, norms are published. He argues that these norms then become static while the population remains dynamic.

According to Erasmus (1997), the developers of the PIB were faced with the following challenges to overcome concerns regarding the standardisation of a test in the development of the PIB:

- Development of generic standards for the population as a whole or for cultural groups.
- Changing standards to be dynamic, like the population.
- Attention to the selection of test items to ensure that the battery is fair to the person, as reflected in the first principle previously discussed.
- Ensuring situation fairness in diverse workplaces.

A situation-specific approach was therefore used in the standardisation of the PIB. This situation-specific approach is also an internationally accepted prerequisite of the APA and the Labour Relations Act, No 66 (1995), of South Africa according to Erasmus (1997). Separate sets of norms, based on culture, were developed as a baseline, but

Erasmus (1997) indicated that situation-specific norms would be developed on an ongoing and dynamic basis.

4.4.2.3 Reliability and validity of the PIB

Limited results regarding the reliability and validity of the PIB were found in the literature review.

Kriel (2001) indicated that the developers of the PIB made their instrument available to users and experts from academic institutions to scrutinise the validity and reliability of the PIB. Schaap (1997) investigated the validity of the PIB for the use in personnel selection and found a high positive correlation between the 16 PF (The Sixteen Personality Factor Questionnaire) and the PIB. He also found a reliability coefficient of 0.93 for the PIB's Index five (Mental Alertness). In Kriel's (2001) research the reliability coefficient ranged between 0.68 and 0.87, with most of the indexes above 0.75. Kriel (2001) reported reliability coefficients for the Mental Alertness sub-test (0.83) and the Vocabulary sub-test (0.76).

The use of a computer scoring system also contributes to the reliability of the PIB.

4.4.2.4 The PIB as part of the selection battery of Technikon Pretoria

Technikon Pretoria currently uses the Potential Index Battery (PIB) as part of its student selection battery. Kriel (2001) pointed out that the main reason for this choice is that it was developed in South Africa and is regarded as culture friendly. She indicated that the Department of Student Counseling was satisfied with the initial investigation regarding the instrument's culture fairness as well as the compliance with relevant regulations and legislation. She pointed out that "This made the instrument acceptable to student bodies

and the Technikon Management, and also from a professional and legal point of view” (Kriel, 2001, p. 102).

The PIB is administered to all students who apply for admission to Technikon Pretoria. The test is administered in a computer laboratory that can accommodate up to 40 students at a time. Qualified staff from the Student Service Bureau administer the test. Scores are automatically calculated at the end of the test and printouts are made. Copies of the PIB results used in this study were obtained from the Student Service Bureau of the Nelspruit campus.

The following cognitive dimensions sub-tests, also known as indexes of the PIB, were used in the 2000 selection battery for students applying for the National Diploma: Human Resources Management at Technikon Pretoria:

- Index three: Reading Comprehension
- Index four: Calculation
- Index five: Mental Alertness

Erasmus and Minnaar (1995) provided the following brief descriptions of these sub-tests in the PIB manual.

- Index three: Reading Comprehension

Ability to read and understand clearly and objectively what the reading matter concerned really conveys.

- Index four: Calculation

Numerical potential referring to the ability to “work” with figures and numbers: to add, subtract and divide.

- Index five: Mental Alertness

The ability to recognise or observe deviations from and differences between seemingly unrelated items and materials. This ability is usually associated with the ability to classify objects correctly and is of special interest in administrative and clerical fields. It is also strongly associated with cognitive ability in general and intellectual “sharpness” in particular.

The PIB also contains sub-tests that measure interpersonal skills. None of these interpersonal skill sub-tests were part of the battery used for student selection in 2000.

4.4.3 The Learning Potential Computerised Adaptive Test (LPCAT)

The LPCAT (Learning Potential Computerised Adaptive Test) was included in this research project as an alternative measure of cognitive ability, resulting from the move in psychometric testing from static measures of cognitive ability to the assessment of learning potential. The LPCAT is a dynamic test administered in a dynamic manner (Murphy, 2002).

The LPCAT was developed between 1993 and 2000 as a measure of learning potential, in response to objections against traditional measures of cognitive ability (De Beer, 2000b). A different form of cognitive assessment was required in South Africa, because differences in educational opportunities and socio-economic backgrounds have an impact on performance in standard cognitive tests. The LPCAT was designed with the aim to assess individuals in the multicultural situation prevailing in South Africa. The LPCAT is a non-verbal, culture-fair measure of the learning potential of an individual in the domain of general reasoning ability (De Beer, 2000c). The LPCAT uses non-verbal items in an attempt to reduce the effect of language on the tests scores (De Beer, 2000b, 2002; Van Eeden et al., 2001).

“Learning potential for the LPCAT is defined as a combination of the pretest performance and the magnitude of the difference between the post-test and pretest scores” (De Beer, 2000c, p. 141). The LPCAT therefore considers the current level of functioning, the potential future level of functioning, and the potential for improving on the present level of performance in the interpretation of the results (De Beer, 2000c). The LPCAT resorts under the measurement approach of dynamic assessment and is based on the principles of Vygotsky’s ZPD theory. There are also some comparisons between the LPCAT and the dynamic assessment measures of Budoff, both utilising non-verbal tasks to measure general abilities of individuals (De Beer, 2000c; Murphy, 2002).

The LPCAT was designed as an instrument that can be used for screening, with the specific objective of preventing unintended discrimination against disadvantaged groups (De Beer, 2000a). Bias analysis, by making use of DIF (Differential Item Functioning), was conducted in the development of the test to eliminate bias in terms of gender, culture, language groups and level of education (De Beer, 2000a, 2000c). It is regarded as a culture-fair test that can be used in a multicultural environment for people from different backgrounds (De Beer, 2000a, 2002).

De Beer (2000c) and Murphy (2002) pointed out the following advantages of using Item response Theory (IRT) and Computerised Adaptive Testing (CAT) in the design of the LPCAT:

- The use of CAT reduces administration time without forfeiting quality or accuracy in measurement.
- Using CAT makes it possible for the test to change for each individual, adapting the degree of difficulty of test items depending on the answers of the testee.
- IRT makes it possible to analyse and select unbiased items in the test development and construction.
- IRT results in more accurate measurement of change scores.

Although it is a computerised test, *no computer literacy is required*. Only the spacebar and the enter key are used in the test. The total testing time is approximately one hour. It can be administered on a personal computer (PC) or a laptop for individuals, or on a computer network for groups (De Beer, 2000b, 2000c). Computer printouts of results are readily available on completion of the test.

The test uses a dynamic test-train-retest format. Learning experiences are provided through standardised computerised training during the assessment process. The LPCAT evaluates the current level of performance, but also determines the extent to which individuals are able to improve their performance when training is provided. In other words, it is also assessing possible future levels of performance. De Beer (2000c) indicated that the pretest measure indicates the present level of performance, and the post-test measure indicates improvement after training. As standard training is provided during the test, examinees' results can be compared. Interactive feedback is provided for practice examples and during training where both the correct answer and relevant feedback are provided (De Beer, 2000c).

4.4.3.1 Administration of the LPCAT

The following information on the administration of the test was obtained from the test manual (De Beer, 2000b). The test takes approximately one hour to administer. There are four sections as part of the assessment procedure:

- Introduction and instructions on computer keys (ENTER and SPACEBAR) to be used during test as well as the answering procedure. Examples are provided before commencing with the pretest.
- The pretest.
- Training and additional examples are provided.

- The post-test is administered.

The LPCAT was administered to the students in the sample at the same time as the PIB was administered to the students. Unfortunately the disks containing the results were damaged and all data were lost. Students were then requested to redo the LPCAT test. Messages and notices were used to inform students to meet in the computer laboratory to redo the LPCAT. However, despite all efforts, only 46 students came to redo the LPCAT test.

4.4.3.2 Standardisation of the LPCAT

A multicultural group consisting of 2450 pupils from 41 schools in Mpumalanga, Northern Cape and the Northern Province was used for the standardisation of the LPCAT. Schools were randomly selected by the HSRC Center for Statistical Support. The standardisation sample further had an almost equal representation of males and females (De Beer, 2000a, 2000c). The percentage distribution of different groups in the standardisation sample was:

TABLE 4.6 DISTRIBUTION OF RACIAL GROUPS IN THE STANDARDISATION SAMPLE OF THE LPCAT (De Beer, 2000c)

Racial group	Percentage representation within the sample
Africans	48.65%
Coloureds	24.49%
Whites	26.86%

De Beer (2000a) pointed out that Indians were excluded from the standardisation sample for two reasons. Firstly, they represent a very small part (less than three percent) of the population of South Africa and secondly, their scores are usually close to those that Whites obtain. The LPCAT meets the requirement that a standardisation sample should be representative of the population for which the test is designed.

4.4.3.3 Reliability and validity of the LPCAT

De Beer (2000c) emphasised the concern raised by Grigorenko and Sternberg (1998) that reliability and validity are often not given sufficient attention in dynamic assessment and she intended to cover this extensively in her design of this dynamic instrument.

The LPCAT's internal consistency reliability indices range between 0.925 and 0.987 (De Beer, 2002).

The use of a large sample, with good representation of various cultural groups in South Africa, contributed to the validity of the LPCAT. Measures were imposed to ensure that test items should not be biased against any racial group. A panel of experts was used to assess the items included in the test. Biased items were identified by means of Differential Item Functioning (DIF) procedures and were removed (De Beer, 2000c; Murphy, 2002).

The computerised format of the test contributes towards validity (Murphy, 2002), because of the accuracy in measurement. The use of standardised training and universal figural items contributes to the face validity of the LPCAT. In the standardisation of the LPCAT, the correlation between the LPCAT and other cognitive tests was investigated to ensure construct validity. For a group of 92 first-year science and technology students, the LPCAT post-test score had a higher correlation with the GSAT than the pretest score. The overall results showed that the two tests measured similar constructs, indicating construct validity (De Beer, 2000a).

In further validation studies of the LPCAT there was a significant correlation with the criterion variables, which provides support for its predictive validity (De Beer, 2003).

The following correlation scores were obtained for the LPCAT post-test for various groups:

TABLE 4.7 CORRELATIONS OF LPCAT POST-TEST SCORES WITH DIFFERENT CRITERIA RESULTS (De Beer, 2003)

Group	Description of criteria	N	r
Adults – low literacy	ABET Literacy Level 1	182	0.44**
	ABET Numeracy Level 2	182	0.49**
	ABET Literacy Level 3	111	0.46**
	ABET Numeracy Level 3	26	0.61**
Grade 8	English proficiency	128	0.61**
	Mathematics proficiency	128	0.67**
	Academic average results	116	0.69**
Grade 12+	Numeracy scores (non-standardised)	146	0.39**
	Language score (non-standardised)	146	0.41**
	2 nd Matriculation Mathematics	158	0.53**
	2 nd Matriculation Science	158	0.45**
	2 nd Matriculation Biology	58	0.31*
	2 nd Matriculation Average	158	0.51**
Technikon 1 st year	Academic average semester 1	89	0.32**
	First year average	89	0.23**
University 4 th year	Academic Average	46	0.05

** statistically significant at $p < 0.01$

r = 0.1 small effect r = 0.3 medium effect r = 0.5 large effect

Table 4.7 summarises the correlation between the LPCAT and criteria for groups with different levels of education. There are statistically significant correlations for all groups, except for fourth-year university students where there was not statistically significant correlation between LPCAT post-test scores and the academic average mark that students obtained. The magnitude of correlations, interpreted as effect size, ranged between medium and large for all groups except for the fourth-year university students (De Beer, 2003). A significant statistical correlation was also found between the LPCAT and the following matric subjects: English, Mathematics and Science, according to De Beer (2000b).

De Beer (2000b, p. 93) concluded that:

For multicultural groups, the LPCAT seems to be a reasonably equitable measure of learning potential within the non-verbal reasoning ability domain.

She further stated:

For academic selection purposes, the best practice would probably be to use learning potential scores together with previous academic results and standard cognitive test results.

De Beer (2000b) suggested further validity investigations of the LPCAT.

4.5 CRITERION MEASURES

A criterion measure is the measure against which the accuracy of the predictors is measured, to determine the predictive validity of the independent variables. Huysamen (1999, p. 137) pointed out that “any variable’s predictive effectiveness can only be as good as the relevant criterion allows it to be”. The choice and reliability of criteria used, are therefore fundamental in a study like this.

A review of other similar research projects in South Africa reflected various ways in which academic performance is used as criterion. Shochet (1986) criticised past research for defining academic success too broadly. He argued that passing an academic year was used, instead of focusing on specific subjects, in determining the predictive value of selection variables. Nunns and Ortlepp (1994) supported Shochet and aimed at finding predictors for individual subjects. Results of the research of Nunns and Ortlepp (1994), however, indicated that predictors of success in one particular subject did not necessarily predict success in other subjects.

Huysamen (1999, 2001), who has done extensive research in student selection in South Africa, indicated that either the MCPM (mean percentage marks) of the first year, or the number of subject credits earned during a period, can be used as criteria. Both of these will be used as criteria in this project.

Internationally, the college grade point average (GPA), calculated as the average performance of all subjects enrolled for, is used as criterion (Brown, 1994; Huysamen, 2001; Wilson, 1980; Zeidner, 1988). Huysamen (2001) and Wilson (1983) refer to the average first-year performance as freshman grade point average (FGPA), which is used as criterion internationally. The MCPM used in South African research is similar to the internationally used GPA.

In international literature, Wilson (1983) also argued that the achievement of students over a three-year period should be used as criterion. Wilson further pointed out that high predictive validity between the selection criteria and first year academic performance, does not necessarily represent a sufficient sample of a student's academic performance for the full duration of a course. A tendency of improved academic performance after the first academic year exists. This is called the late-blooming hypothesis (Huysamen, 2000; Wilson, 1983). Wilson (1983) explains the late-blooming hypothesis as the tendency that students from minority groups take longer to adjust to tertiary education, and that those students, after a year of "academic acculturation", show improvement in their

performance. As opportunities and resources are scarce, they should not be wasted on individuals who do not have the potential to obtain a qualification (Jackson & Young, 1988; Maxwell, 1996; Nunns & Ortlepp, 1994; Rutherford & Watson, 1990; Uys, 1994; Zaaïman, 1998).

As indicated earlier, in the definition of Jonker (1994), Louw (1993, 1996), as well as Uys (1994), they indicated that the selection process should identify applicants with the necessary potential to *complete courses successfully*. Louw (1996) also concluded that academic performance throughout the three years of study should be used as a criterion. The emphasis, according to these authors, is clearly on the completion of a course.

Huysamen (1999) pointed out that reliability of criteria could be affected if students register for different combinations of subjects, or even if different lecturers mark the same combination of subjects. This can then affect the correlations with other variables.

The sample of students used in this study is from one course, which offers only one optional subject, and thirteen compulsory subjects. The same lecturer marks the same tests and assignments of all students. There is thus very little effect on the reliability of students' marks, that could impact on the correlation calculations in this study.

There are primary and secondary criteria that will be used in this project.

- Primary criterion

The primary criterion that will be used is the MCPM (mean of percentage mark), in each of the three years. As indicated earlier, MCPM are similar to the GPA used internationally. The actual average academic performance of all students over a three-year period, in other words the MCPM, will be calculated per year. This will be used as the primary criterion against which the predictive validity of matric results, the PIB and the LPCAT will be measured.

- Secondary criterion

The number of subjects that a student passes over the three-year period, as percentage of the required fourteen subjects a student should pass to complete the National Diploma, will be used as secondary criterion. The rationale to include this criterion is based on the definitions of student selection as highlighted in section 3.2.1.

4.6 DATA GATHERING AND PROCESSING

Data regarding gender, tertiary academic scores, matric results, bridging course status, PIB scores and LPCAT scores were captured on a spreadsheet.

Data regarding the tertiary academic performance of students were obtained through computer printouts from the Academic Administration Department of the Nelspruit satellite campus of Technikon Pretoria. The actual final marks of students in each subject were captured. Final marks are the average between the mark obtained during a semester (called a predicate mark) and the mark obtained in the examination. There are fourteen subjects in the National Diploma: Human Resources Management. The average percentage of each student's academic performance per semester was calculated as well as the average per year for each of the three years.

The number of semester subjects which students passed per year was determined and captured for each year of study in the three-year period. The total number of subjects passed over the three years was then calculated as percentage of the fourteen subjects that a student should have passed over a three-year period. This data will be used to determine a student's progress towards completion of the National Diploma.

The following biographical data were also obtained from the files of students. Numerical codes were used in capturing data with non-numerical values.

- Gender (male/female)
- Bridging course status (yes/no)

The gathering of the data of the PIB and the LPCAT was previously described in section 4.4.2 and 4.4.3. The actual scores obtained in the two psychometric tests were captured on the spreadsheet.

All statistics were carried out using the SAS statistical package. The findings will be reported in the following chapter.

4.7 CHAPTER SUMMARY

In this chapter the research sample and research procedures applied in this project were described. The measuring instruments that were used were also discussed and reference was made to the criteria used in the correlation study.

In the next chapter the results of correlation and multiple regression analysis calculations will be reported. The results will then be interpreted and discussed.

CHAPTER 5

RESEARCH RESULTS

5.1 INTRODUCTION

In this chapter the results of the present study will be reported. The aim of this study is to determine which of the variables used can significantly predict tertiary academic performance. The statistical and regression analysis of data makes it possible to report on the predictive validity of various variables that are, or could be, used in a selection battery for the National Diploma: Human Resources Management at Technikon Pretoria Nelspruit satellite campus.

The variables that were used, are:

- Matric results, in terms of MSPT (Matric symbol point total) and certain matric subjects
- PIB
- LPCAT
- Bridging course
- Gender

SAS, a statistical software system, was used to analyse data that were captured in an Excel spreadsheet. This chapter contains statistical results which will be reported in the following order:

- Descriptive statistics of the independent variables and criterion measures.
- Correlation coefficients between independent variables and criterion measures.
- *t*-Tests to determine differences between groups.
- Regression analysis results.

The statistical values obtained will be reported in full, according to the objectives set in chapter one. The results obtained in this project will be interpreted in accordance with the literature review and other similar research projects reported on in chapter three.

5.2 DESCRIPTIVE STATISTICS

Descriptive statistics provide an overview of the data that were obtained in the research project. The descriptive statistics of some of the variables are presented in tables 5.1 and 5.2. Descriptive statistics of certain variables, such as gender and bridging course, as well as whether students did commercial subjects at school or not, cannot be determined as these variables does not have any numerical values. These tables reflect information regarding the sample size (n) for each variable, the mean, the standard deviation as well as the minimum and maximum values.

The mean reflects the average score of the sample and the standard deviation is a statistical measure, which reflects how data deviate from the mean (Tredoux & Durrheim, 2002; Wegner, 1993). The minimum and maximum values reflect the extreme values within the sample.

5.2.1 Descriptive statistics of the independent variables (predictors)

The independent variables' descriptive statistics will be provided in table 5.1. These variables are the two psychometric tests, namely the PIB and LPCAT, and the matrix results.

TABLE 5.1 DESCRIPTIVE STATISTICS OF THE INDEPENDENT VARIABLES

Variable	N	Mean	Std dev	Minimum	Maximum
LPCAT pretest score (80)	46	47.52	6.97	34.45	63.64
LPCAT post-test score (80)	46	47.63	7.13	34.31	62.18
PIB Reading Comprehension (5)	63	2.67	0.84	1	5
PIB Calculations (5)	63	1.29	0.52	1	3
PIB Mental Alertness (5)	64	1.95	0.68	1	4
Matric symbol point total (MSPT) (40)	70	19.1	3.60	11	27
Matric English (8)	70	4.4	1.11	1	7

(The possible maximum score is indicated in the variable column in brackets.)

As indicated earlier, the total sample consisted of 72 students, but due to the unforeseen loss of data, there were only LPCAT scores available for 46 students within the sample. The number of respondents of the PIB sub-tests also varies due to students who get automatic access to the National Diploma on completion of a bridging course, and are exempted from the PIB as selection criteria to gain admission to the National Diploma.

The standard deviation for the LPCAT scores indicated a reasonable distribution of scores. The mean scores are somewhat lower than what is typically found for first-year Technikon students (De Beer, 2000b, 2003). The descriptive statistics of the PIB reflected that most of the students performed the best on the Reading Comprehension sub-test of the PIB. The students further obtained very similar mean values in the Mental Alertness and Calculation sub-tests. All three these variables reflected a normal distribution of data.

The MSPT's (Matric symbol point total) standard deviation is relatively small, indicating that the data are closely distributed around the mean. The mean score is 19.1 out of a possible 40, which is equal to 47.75 percent. This represents the average of the students' matric symbol point total.

Descriptive statistics of the mark that students obtained in matric English, are also provided. The performance of the students in a single subject is also investigated, to determine the possible relationship between English proficiency and academic performance.

5.2.2 Descriptive statistics of the criterion

There are primary and secondary criteria that will be used in this project.

5.2.2.1 Primary criterion

The primary criterion that will be used is the mean of percentage marks, MCPM, in each of the three years. MCPM is similar to the GPA used internationally.

The following acronyms will be used:

MCPM 1	-	Mean of percentage marks in the first year
MCPM 2	-	Mean of percentage marks in the second year
MCPM 3	-	Mean of percentage marks in the third year
MCPM 1 A	-	Mean of percentage marks - first semester in the first year
MCPM 1 B	-	Mean of percentage marks - second semester in the first year
MCPM 2 A	-	Mean of percentage marks - first semester in the second year
MCPM 2 B	-	Mean of percentage marks - second semester in the second year
MCPM 3 A	-	Mean of percentage marks - first semester in the third year
MCPM 3 B	-	Mean of percentage marks - second semester in the third year

A and B used above represent the first semester (A) and the second semester (B) of the respective years.

5.2.2.2 Secondary criterion

The number of subjects that a student passes, as percentage of all the subjects a student should pass to obtain the diploma, over the three-year period, will be used as secondary criterion in the project. The reason for including this as secondary criteria is a result of the focus of the definition of student selection on completing a course, as discussed in section 3.2.1.

TABLE 5.2 DESCRIPTIVE STATISTICS OF THE CRITERION VARIABLES

Variable	N	Mean	Std dev	Minimum	Maximum
MCPM 1 (100)	70	46.2	7.02	29	65
MCPM 2 (100)	56	51.96	11.00	0	73
MCPM 3 (100)	45	55.47	7.62	31	74
MCPM 1 A (100)	70	45.12	7.03	29	64
MCPM 1 B (100)	61	48.16	7.04	34	66
MCPM 2 A (100)	56	50.84	10.69	0	67
MCPM 2 B (100)	46	55.91	8.69	27	80
MCPM 3 A (100)	45	54.84	6.93	35	68
MCPM 3 B (100)	42	55.62	9.69	27	79
Percentage subjects passed over three years (100)	70	40.41	33.96	0	100
Difference between MCPM 1B – MCPM 1A	61	1.87	6.14	-14	15
Difference between MCPM 2B – MCPM 2A	46	2.82	2.28	-11	17
Difference between MCPM 3B – MCPM 3A	41	1.10	1.10	-19	15
Difference between MCPM 2 – MCPM 1	56	4.04	4.04	-46	22
Difference between MCPM 3 – MCPM 2	43	0.72	0.72	-14	15

(The possible maximum score is indicated in the variable column in brackets.)

In table 5.2 it is noticeable that the number of students per academic year and semester varies as a result of students temporarily or permanently ending their studies. Although some students do progress from their first to second and third academic year, they might be repeating failed subjects. This implies that, although students may be in their third

year of study, they might still be far from completing the fourteen subjects to qualify for their National Diploma: Human Resources Management. Figure 5.1 reflects that the drop out from the first year first semester to the second semester of the second year is relatively bigger than the drop-out from the second year's second semester to third year. This can be viewed as a result of the adjustment to tertiary education. There is an indication that the adjustment influences students more in their first year and the first semester of their second year, than later in their studies.

The standard deviation for the number of subjects passed as percentage of total subjects is very high.

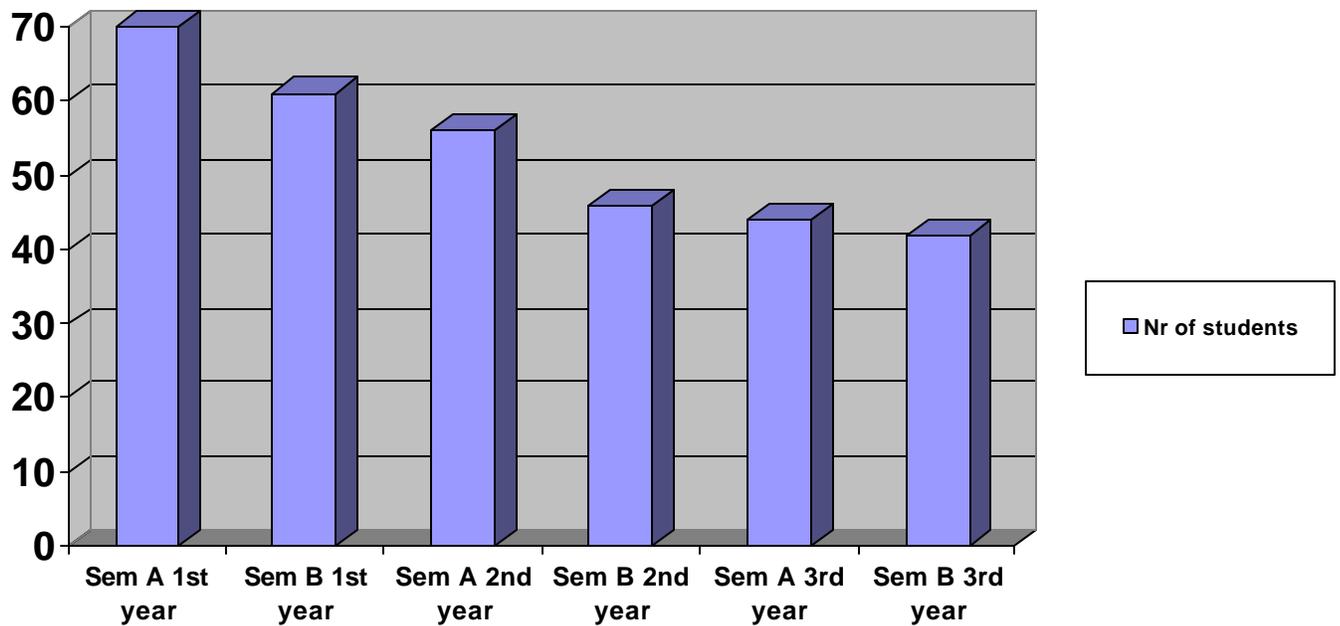


Figure 5.1 Number of students per semester over a three-year period

The standard deviations of MCPM 1 to MCPM 3 are relatively high, reflecting a wider distribution of data resulting in a flatter distribution curve. There was one extremely low value, zero, in the first semester on second year level. An investigation of the data reflected that one student registered for three subjects in the first semester of her second

year, but did not obtain the minimum of 40 percent in any one of the three subjects, and therefore did not qualify to write the examination in any of those subjects. As the final marks (the average between the semester mark and examination mark per subject) were used to calculate annual MCPM's, this particular student had no final mark and therefore obtained zero as MCPM. As this is an outlier in the data, it might affect the statistical results. These data were, however, retained to use in earlier semesters' MCPM's and years' MCPM's.

There is also an indication that the MCPM of students improved progressively from their first to second year, with a further improvement in the third year. Therefore the late-blooming hypothesis that Huysamen (2000) and Wilson (1983) refer to, is also evident in this research project.

The semester MCPM's were also calculated. The average mark of semester A and B in the first year does not even reach 50 percent, which is the required mark to pass. The MCPM from the second year onwards represents an average above 50 percent. From the second semester of the second year it appears as if students' academic performance stabilizes and reaches a plateau and their academic performance remains on very similar levels thereafter, ranging between 54.48 and 55.91 percent (see figure 5.2).

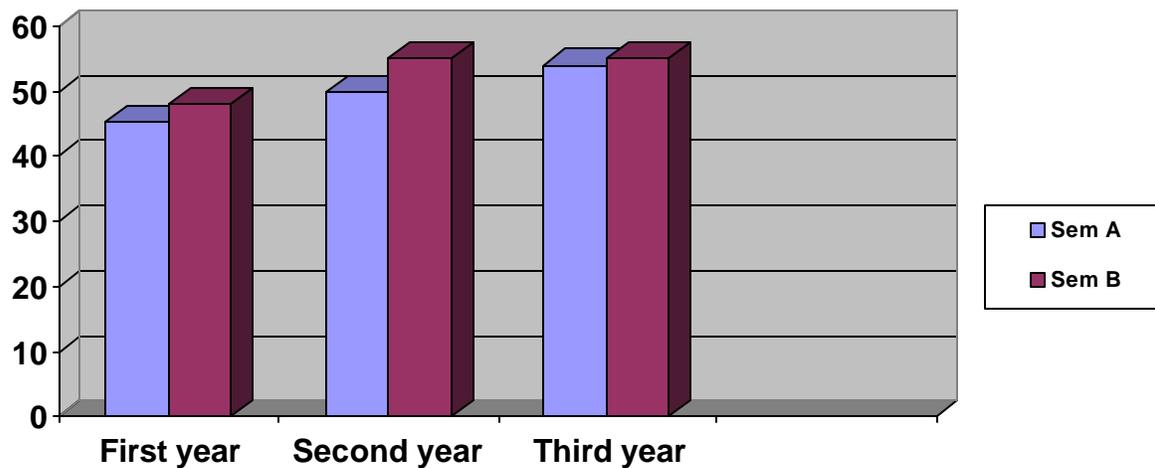


Figure 5.2 MCPM per semester for the first, second and third year

The histogram in figure 5.2 represents the descriptive statistics of the data in table 5.2. This figure reflects the increase in the students' average performance over the six semesters in the three years of study. There is an increase between the three first semesters and similar performances between the three last semesters, also reflecting the late-blooming hypothesis and the plateau which students' performance shows once they adjust to tertiary education.

TABLE: 5.3 *t*-TESTS RESULTS REFLECTING THE DIFFERENCE IN ACADEMIC PERFORMANCE BETWEEN SEMESTERS AND YEARS

Variable	N	<i>d</i>	T Value	Pr > t
Difference between MCPM 1B – MCPM 1A	61	0.42	2.38	0.02 *
Difference between MCPM 2B – MCPM 2A	46	0.52	2.49	0.02 *
Difference between MCPM 3B – MCPM 3A	41	0.09	1.01	0.32
Difference between MCPM 2 – MCPM 1	56	0.64	3.22	0.00 **
Difference between MCPM 3 – MCPM 2	43	0.38	0.84	0.45

*significant at $p < 0.05$ d = 0.2 small effect	** significant at $p < 0.01$ d = 0.5 medium effect	*** significant at $p < 0.001$ d = 0.8 large effect
--	---	--

Table 5.3 reflects the values that were obtained from *t*-tests to determine the difference in academic performance between the first and second semester for the first to the third year MCPM's. The difference in academic performance between the first and second semester of each year, the difference between the second and first year, as well as the difference between the third and second year, were calculated.

There is a statistically significant ($p = 0.02$) difference between the second semester (B) and first semester (A) of the first year. The effect size for this difference is almost medium ($d = 0.42$). The effect size for the difference between the second semester (B) and first semester (A) of the second year is medium ($d = 0.52$), at a statistically

significant level ($p = 0.02$). There are no statistically or practically significant differences between the second and first semester of the third year, reflecting the plateau that student's academic performance reach.

The following results were obtained in terms of the academic performances in different years of study. The effect size for the difference between the second and first year mean of percentage marks is slightly higher than medium ($d = 0.64$). This difference is also statistically highly significant. There are no statistically significant differences between the third and second year MCPM's. In terms of effect size, the difference is between small and medium.

These *t*-tests and effect size calculations confirm the conclusions, drawn from the descriptive statistics, that students' performances stabilise during the second year and that academic performance stays constant in the third year.

5.3 CORRELATION RESULTS

The aim of the study was to determine the predictive validity of predictor variables that can be used in student selection. The predictor variables listed in table 4.2 were used for the correlation analysis. Correlation results will be interpreted in terms of statistical significance (*p*-value) and practical significance (effect size).

5.3.1 Correlation between predictors

The various predictors used in this study were subjected to the correlation analysis to determine any statistically significant correlations between the various predictors used in the study.

TABLE 5.4 PEARSON CORRELATION COEFFICIENT BETWEEN PREDICTORS

	LPCAT pretest			LPCAT post-test			MSPT		
	N	r	p	N	r	p	N	r	p
LPCAT pretest				46	0.91	0.00***	46	0.17	0.27
LPCAT post-test	46	0.91	0.00***				46	0.36	0.01**
PIB Reading Comprehension	40	0.24	0.13	40	0.20	0.21	63	0.10	0.42
PIB Calculations	41	-0.04	0.81	41	-0.10	0.54	63	-0.04	0.76
PIB Mental Alertness	41	0.24	0.11	41	0.34	0.03*	64	0.19	0.14

*significant at $p < 0.05$

$r = 0.1$ small effect

** significant at $p < 0.01$

$r = 0.3$ medium effect

*** significant at $p < 0.001$

$r = 0.5$ large effect

In table 5.4 the LPCAT pretest and post-test correlated highly ($r = 0.91$) with one another, at a statistically significant level of $p < 0.001$. The magnitude of this correlation, in terms of effect size, is very large. This reflects multi-colinearity between these two variables. Multi-colinearity indicates that the predictor variables are strongly inter-correlated (Tredoux & Durrheim, 2002).

The correlation results in table 5.4 indicate a statistically significant correlation between the LPCAT post-test and the MSPT ($r = 0.36$; $p < 0.01$). The magnitude of this correlation effect size, in terms of effect size, can be considered medium. The LPCAT post-test also correlated significantly with the PIB Mental Alertness sub-test ($r = 0.34$; $p < 0.05$). The correlation effect size is medium. None of the other predictors correlate statistically significantly with one another.

TABLE 5.5 SPEARMAN CORRELATION COEFFICIENT BETWEEN PREDICTORS

	LPCAT pretest			LPCAT post-test			MSPT		
	N	r	p	N	r	p	N	r	p
LPCAT pretest				46	0.90	0.00***	46	0.20	0.17
LPCAT post-test	46	0.90	0.00***				46	0.40	0.01**
PIB Reading Comprehension	40	0.24	0.14	40	0.20	0.21	63	0.09	0.48
PIB Calculations	41	0.04	0.81	41	-0.08	0.62	63	-0.05	0.70
PIB Mental Alertness	41	0.21	0.19	41	0.30	0.06	64	0.24	0.06

*significant at $p < 0.05$	** significant at $p < 0.01$	*** significant at $p < 0.001$
$r = 0.1$ small effect	$r = 0.3$ medium effect	$r = 0.5$ large effect

In table 5.5 the LPCAT pretest and post-test correlated strongly with each other at a statistically significant level ($r_s = 0.9$; $p < 0.001$). The magnitude of this correlation in terms of effect size is very large. The Spearman correlation coefficient further indicates a statistically significant correlation between the LPCAT post-test and the MSPT ($r_s = 0.40$; $p < 0.01$). The correlation effect size is between medium and large.

5.3.2 Correlation between predictors and criteria

The magnitude and statistical significance of correlations between the predictors listed in table 4.2 and the different criteria will indicate whether the predictors can be considered valid predictors of future academic performance.

Huysamen (1999) pointed out that, because students need a minimum predicate mark to be admitted to write examinations, it can decrease the correlation between the predictor and criterion, as the range of scores are restricted by certain students being excluded from

the examinations. This should be considered in interpreting the correlation values that follow.

5.3.2.1 Correlation between matric results and MCPM from first year to third year

It was an objective of this research project to determine the correlation of matric results (MSPT) and tertiary academic performance, as well as the relationship between specific matric subjects and academic performance during the three years.

There are slight differences between the Pearson and Spearman correlation coefficients in table 5.10. This can be ascribed to differences in calculating these coefficients as described in section 4.2.

a. Correlation between MSPT and MCPM from first year to third year

In calculating the correlation between matric results and tertiary academic performance, the mean percentage mark (MCPM) and the matric symbol point total (MSPT) were used.

TABLE 5.6 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF MSPT AND MCPM 1 TO MCPM 3

	N	Pearson		Spearman	
		MSPT	MSPT	MSPT	MSPT
		r	p	r _s	p
MCPM 1	70	0.23	0.05*	0.25	0.04*
MCPM 2	56	0.13	0.34	0.20	0.14
MCPM 3	45	0.31	0.04*	0.33	0.03*
*significant at $p \leq 0.05$		** significant at $p < 0.01$		*** significant at $p < 0.001$	
r = 0.1 small effect		r = 0.3 medium effect		r = 0.5 large effect	

The actual probability value of the Pearson correlation coefficient in table 5.6 is $p = 0.0519$ and was rounded off to 0.05. This probability value is a borderline case of statistical significance. Since the Spearman correlation coefficient is regarded as more meaningful (as explained in section 4.2) than the Pearson correlation coefficient, the correlation between MSPT and MCPM 1 ($r_s = 0.25$; $p < 0.05$) is regarded as statically significant. The magnitude of this correlation, in terms of effect size, for the Spearman correlation coefficient, is between small and medium. Both the Pearson and Spearman correlation coefficients for MSPT and MCPM 2 are not statistically significant. The Pearson correlation coefficient ($r = 0.31$; $p < 0.05$) and the Spearman correlation coefficient ($r_s = 0.33$; $p < 0.05$) are statistically significant for the correlation between MSPT and MCPM 3. The effect size of this correlation value can be considered medium.

As all students in the sample came from a DET (Department of Education and Training) school, the findings of relatively low correlations of tertiary academic performance on first-year level with MSPT, as well as no statistically significant correlations between MSPT and MCPM 2, are in accordance with the research findings of Miller (1992), Shochet (1986), Skuy et al. (1996) and Zaaiman (1998).

TABLE 5.7 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF MSPT AND THE PERCENTAGE OF SUBJECTS PASSED OVER THE THREE-YEAR PERIOD

	N	Pearson MSPT		Spearman MSPT	
		r	p	r_s	p
Percentage of subjects passed over the three-year period	70	0.23	0.05*	0.27	0.02*
*significant at $p \leq 0.05$		** significant at $p < 0.01$		*** significant at $p < 0.001$	
r = 0.1 small effect		r = 0.3 medium effect		r = 0.5 large effect	

The actual probability value of the Pearson correlation coefficient in table 5.7 was $p = 0.0511$ which was rounded off to 0.05. This probability value is a borderline case of statistical significance. Since the Spearman correlation coefficient is regarded as more meaningful (as explained in section 4.2) than the Pearson correlation coefficient and the Spearman correlation coefficient between MSPT and MCPM 1 is $r_s = 0.27$ with $p = 0.02$, the correlation is regarded as statically significant. The magnitude of the correlation, in terms of effect size, is between small and medium.

b. Correlation between matric English and MCPM from first year to third year

The correlation between matric English and tertiary academic performance was investigated to determine whether a student’s English proficiency has an influence on his or her tertiary academic performance. The Swedish formula (section 4.4.1) was used to allocate a numeric value to the symbol that students obtained for matric English.

TABLE 5.8 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF MATRIC ENGLISH AND MCPM 1 TO MCPM 3

	N	Pearson		Spearman	
		Matric English	Matric English	Matric English	Matric English
		r	p	r_s	p
MCPM 1	70	0.03	0.84	0.03	0.79
MCPM 2	56	0.04	0.79	- 0.01	0.96
MCPM 3	45	0.05	0.75	0.04	0.78

*significant at $p < 0.05$

$r = 0.1$ small effect

** significant at $p < 0.01$

$r = 0.3$ medium effect

*** significant at $p < 0.001$

$r = 0.5$ large effect

The Pearson and Spearman correlation coefficients between the performance of a student in matric English and MCPM 1 to MCPM 3 reflected almost no correlation and were not statistically significant.

TABLE 5.9 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF MATRIC ENGLISH AND THE PERCENTAGE OF SUBJECTS PASSED OVER THE THREE-YEAR PERIOD

	N	Pearson Matric English		Spearman Matric English	
		r	p	r _s	p
Percentage of subjects passed over the three-year period	70	0.03	0.84	0.05	0.66
*significant at p < 0.05 r = 0.1 small effect		** significant at p < 0.01 r = 0.3 medium effect		*** significant at p < 0.001 r = 0.5 large effect	

The correlation values for both the Pearson and Spearman correlation coefficients for the relationship between matric English and the secondary criterion of this research project, namely, the percentage of subjects passed over the three-year period, reflected almost no correlation and were not statistically significant.

These findings are contrary to the view that English proficiency can have an impact on academic performance (De Beer, 2000c; Huysamen, 1999, 2000). A question can, however, be raised about the validity of matric English as a measure of English proficiency.

Further calculations were done to determine whether other matric subjects, like commercial subjects, could have an impact on students' academic performance. This will be reported on in section 5.4.

5.3.2.2 Correlation between LPCAT scores and MCPM from first to third year

To be able to report on the correlation between the academic performance of students and a dynamic psychometric instrument, the following correlation coefficients were calculated.

TABLE 5.10 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF THE LPCAT PRETEST AND POST-TEST SCORES WITH MCPM 1 TO MCPM 3

Predictor	N	Pearson				Spearman			
		LPCAT pretest		LPCAT post-test		LPCAT pretest		LPCAT post-test	
		r	p	r	p	r	p	r	p
MCPM 1	46	0.31	0.04 *	0.29	0.05*	0.23	0.12	0.25	0.09
MCPM 2	35	0.50	0.00**	0.46	0.00**	0.44	0.00**	0.40	0.02*
MCPM 3	28	0.54	0.00**	0.54	0.00**	0.55	0.00**	0.54	0.00**

*significant at $p \leq 0.05$
 $r = 0.1$ small effect

** significant at $p < 0.01$
 $r = 0.3$ medium effect

*** significant at $p < 0.001$
 $r = 0.5$ large effect

The Pearson and Spearman correlation results of table 5.10 between the LPCAT pretest and the MCPM's will first be discussed. Thereafter will the Pearson and Spearman correlation results, reflecting the relationship between the LPCAT post-test and the MCPM's be discussed.

As indicated in section 5.3.1, the LPCAT pretest and post-test correlate highly with each other at a statistically significant level ($r = 91$; $p < 0.0001$). There are therefore not significant differences between the LPCAT pretest's and post-test's correlations with the MCPM's during the three years.

The Pearson correlation coefficient in table 5.10, reflecting the relationship between LPCAT pretest and MCPM 1, was statistically significant ($r = 0.31$; $p < 0.05$). The practical significance and magnitude of this correlation, in terms of effect size, is medium. The Pearson correlation coefficient for the LPCAT pretest with MCPM 2 ($r = 0.50$; $p < 0.01$) and MCPM 3 ($r = 0.54$; $p < 0.01$) were both statistically significant. The effect size of these correlation values reflects a large practical significance for both these correlation values.

The Spearman correlation coefficient also reflected a statistically significant correlation between the LPCAT pretest with MCPM 2 ($r_s = 0.44$; $p < 0.01$) and MCPM 3 ($r_s = 0.55$; $p < 0.01$), with the effect size of these correlation ranging between medium and large. The correlation on first-year level ($r = 0.23$; $p = 0.12$) was not statistically significant.

The difference between the Pearson and Spearman correlation coefficient, reflecting the relationship between MCPM 1 and the LPCAT pretest, is probably a result of the rank order of the data, which can be a result of the wider distribution of the metric data as reflected in the descriptive statistics in table 5.2.

The Pearson correlation coefficient results reflected a statistically significant correlation ($r = 0.29$; $p < 0.05$) between the LPCAT post-test and MCPM 1. In this instance the actual p-value was $p = 0.0484$ and was rounded of to 0.05. Therefore the correlation value of $r = 0.29$ is regarded as statistically significant. The practical significance of this correlation in terms of effect size is medium. The Spearman correlation coefficients did not reflect a statistically significant relationship between the LPCAT post-test and MCPM 1.

The Pearson correlation coefficient in table 5.10 further reflected a positive statistically significant correlation between the LPCAT post-test with MCPM 2 ($r = 0.46$; $p < 0.01$) and MCPM 3 ($r = 0.54$; $p < 0.01$). This reflects a large correlation effect size. The Spearman correlation coefficient between the LPCAT post-test and the second year ($r_s = 0.40$; $p < 0.05$) and third year ($r_s = 0.54$; $p < 0.01$) results were statistically significant.

The correlation effect size of the second and third year averages with LPCAT post-test is medium and large, respectively.

The Pearson correlation coefficients of the LPCAT pretest with the first ($r = 0.31$), second ($r = 0.50$) and third year ($r = 0.54$) level's MCPM results, are slightly higher than the correlation coefficients of the LPCAT post-test with the MCPM results of the first ($r = 0.29$), second ($r = 0.46$) and third-year ($r = 0.54$) level. All these correlations are statistically significant as reflected in table 5.10.

TABLE 5.11 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF THE LPCAT PRETEST AND POST-TEST SCORES WITH THE PERCENTAGE OF SUBJECTS PASSED OVER THE THREE-YEAR PERIOD

Predictor	N	Pearson				Spearman			
		LPCAT Pretest		LPCAT Post-test		LPCAT Pretest		LPCAT Post-test	
		r	p	r	p	r	p	r	p
Percentage subjects passed over the three-year period	46	0.41	0.00**	0.39	0.00**	0.37	0.01*	0.39	0.00**

*significant at $p < 0.05$

$r = 0.1$ small effect

** significant at $p < 0.01$

$r = 0.3$ medium effect

*** significant at $p < 0.001$

$r = 0.5$ large effect

The Pearson and Spearman correlation values in table 5.11, reflecting the strength of the relationship between the percentage subjects passed over the three-year period with the LPCAT pretest and post-test, reflected statistically significant correlations ($p < 0.01$ and p

< 0.001) with a r-value that ranged between $r = 0.41$ and $r_s = 0.37$. The magnitude of these correlations, interpreted in terms of effect size, are medium.

5.3.2.3 Correlation between sub-tests of the PIB and MCPM from first to third year

The project also aimed at determining the correlation coefficients between the PIB and tertiary academic performance over the three-year period.

TABLE 5.12 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF THE PIB SUB-TEST SCORES WITH MCPM 1 TO MCPM 3

		MCPM 1 (Pearson)	MCPM 1 (Spearman)	MCPM 2 (Pearson)	MCPM 2 (Spearman)	MCPM 3 (Pearson)	MCPM 3 (Spearman)
PIB Reading Comprehension	r	0.12	0.13	0.18	0.02	0.05	0.00
	p	0.33	0.32	0.21	0.89	0.75	0.98
	N	63	63	49	49	41	41
PIB Calculation	r	-0.08	-0.10	-0.05	-0.17	-0.10	-0.05
	p	0.52	0.42	0.73	0.25	0.52	0.74
	N	63	63	49	49	42	42
PIB Mental Alertness	r	0.07	0.00	0.04	-0.06	0.03	-0.11
	p	0.59	0.97	0.78	0.68	0.86	0.49
	N	64	64	50	50	42	42

*significant at $p < 0.05$

$r = 0.1$ small effect

** significant at $p < 0.01$

$r = 0.3$ medium effect

*** significant at $p < 0.001$

$r = 0.5$ large effect

Table 5.12 reflected very low and not statistically significant correlations between PIB Reading Comprehension and MCPM 1 to MCPM 3 for both the Spearman and Pearson correlation.

The Spearman and Pearson correlation calculations reflected a low negative relationship between the PIB Calculation sub-test and MCPM 1 to MCPM 3 which was not statistically significant.

All the Pearson correlation coefficients between PIB Mental Alertness and the academic performance of students over the three-year period were very low and not statistically significant. In this case, all the Spearman correlations reflected a negative correlation, which was statistically insignificant.

TABLE 5.13 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF THE PIB SUB-TESTS WITH THE PERCENTAGE OF SUBJECTS PASSED OVER THE THREE-YEAR PERIOD

	PIB READING COMPREHENSION				PIB CALCULATION				PIB MENTAL ALERTNESS			
	Pearson		Spearman		Pearson		Spearman		Pearson		Spearman	
	r	p	r	p	r	p	r	p	r	p	r	p
Percentage subjects passed over the three-year period	0.11	0.38	0.13	0.30	-0.08	0.52	-0.06	0.62	0.01	0.93	-0.04	0.78
<p>*significant at $p < 0.05$ ** significant at $p < 0.01$ *** significant at $p < 0.001$ $r = 0.1$ small effect $r = 0.3$ medium effect $r = 0.5$ large effect N = 63</p>												

Table 5.13 reflects that there is fairly low r-values that is not statistically significant between any of the sub-tests of the PIB and the secondary criterion of this research project, namely the number of subjects that students passed as percentage of the total subjects that a student should pass over a three-year period.

5.3.3 Correlation between academic performance of successive years

There is evidence in the literature review, reported in section 3.6.4, of large and significant correlations between first year performance and the performance in successive years of study. The data of this project were also tested for this phenomenon.

TABLE 5.14 PEARSON AND SPEARMAN CORRELATION COEFFICIENTS OF THE MCPM's OF EACH YEAR WITH SUCCESSIVE YEARS

		MCPM 1 (Pearson)	MCPM 1 (Spearman)	MCPM 2 (Pearson)	MCPM 2 (Spearman)
MCPM 2	r	0.53	0.62		
	p	0.00***	0.00 ***		
	N	56	56		
MCPM 3	r	0.51	0.41	0.71	0.65
	p	0.00 ***	0.00 **	0.00 ***	0.00 ***
	N	45	45	43	43

*significant at $p < 0.05$

** significant at $p < 0.01$

*** significant at $p < 0.001$

$r = 0.1$ small effect

$r = 0.3$ medium effect

$r = 0.5$ large effect

The researcher noted a difference between the number of students included in the correlation analysis between first and third year ($N = 45$), as well as between second and third year ($N = 43$). An investigation into the data revealed that one student did not return for the second semester of the second year and another student did not qualify to write exam during the first semester of the second year. Both these students returned for a third year of studies, making a comparison between 45 students' performance on first and third year possible.

The Pearson correlation results in table 5.14 indicate a statistically significant correlation between MCPM 1 and MCPM 2 ($r = 0.53$; $p < 0.001$). The correlation between MCPM 2 and MCPM 3 ($r = 0.71$; $p < 0.001$) is also statistically significant as well as the correlation between MCPM 1 and MCPM 3 ($r = 0.51$; $p < 0.001$). The magnitude of all these correlations values, interpreted in terms of effect size, is large.

The Spearman correlation coefficients were also statistically significant for the correlation between MCPM 1 and MCPM 2 ($r_s = 0.62$; $p < 0.001$) as well as between MCPM 2 and MCPM 3 ($r_s = 0.65$; $p < 0.001$). MCPM 1 and MCPM 3 also correlated statistically significant ($r_s = 0.41$; $p < 0.01$). The magnitude of these correlations, in terms of effect size, is between medium and large.

These findings are in accordance with those of Ayaya (1996), Huysamen (2000) and Jawitz (1995). These results support the late-blooming hypothesis that indicate that first year performance is a better predictor of subsequent academic performance than matric results.

5.4 *t*-TESTS TO DETERMINE THE DIFFERENCE BETWEEN GROUPS

In addition to correlation analysis, *t*-tests were used as alternative to determine the statistical significance of the differences between the means of two independent groups' academic performance, in terms of their MCPM's per academic year. The effect size (*d*) of the differences was also calculated in each instance to be able to interpret and report on the practical significance of the differences.

The following biographical data were used to investigate the difference in the academic performance between groups in the following three categories:

- Gender
- Bridging course status
- Commercial subjects at school status

A test of normality was conducted on the data used in the *t*-test calculations for the two gender groups. It reflected to be normal and the interpretation of the data that follow is therefore regarded as reliable.

TABLE 5.15 DIFFERENCE BETWEEN MCPM SCORES OF GENDER GROUPS OVER A THREE-YEAR PERIOD

Academic year	F	Pr > F	d	N	Males		Females		
					Mean	Std Dev	N	Mean	Std dev
First year	2.55	0.12	0.39*	24	47.67	7.59	45	45.00	6.03
Second year	5.68	0.02*	0.71**	20	55.95	7.74	35	49.09	9.60
Third year	1.2	0.28	0.34*	17	56.62	6.67	27	54.11	7.04

* d = 0.2 small effect

** d = 0.5 medium effect

*** d = 0.8 large effect

The effect size of the difference between males (MCPM 2 = 55.95%) and females (MCPM 2 = 49.09%) academic performance in the second year, reflected an above medium to large practical significance (d = 0.71), which was also statistically significant (p = 0.02). There were not statistical or practical significant differences on first and third year level.

The second category, where the difference between two groups' performances was calculated, is the bridging course status of students.

TABLE 5.16 DIFFERENCE BETWEEN MCPM SCORES OF GROUPS, BASED ON BRIDGING COURSE STATUS OVER A THREE-YEAR PERIOD

Academic year	F	Pr > F	d	N	Bridging course		No bridging course		
					Mean	Std dev	N	Mean	Std dev
First year	0.70	0.41	0.21*	26	47.12	6.80	44	45.65	7.17
Second year	0.19	0.66	0.13*	22	52.77	7.30	34	51.44	12.92
Third year	0.71	0.40	0.26*	17	54.24	8.04	28	56.21	7.04

* d = 0.2 small effect

** d = 0.5 medium effect

*** d = 0.8 large effect

The *t*-test reflected a small ($d = 0.21$; $d = 0.13$ and $d = 0.26$) and not statistically significant difference between the MCPM's (over the three-year period) of students who did a bridging course and those who did no bridging course. It therefore appears as if the bridging course status of a student has very little impact on academic performance. The test for normality for the MCPM 3 comparison reflected that the data used were not normally distributed and therefore these scores should be used with caution. The other data samples for the MCPM 1 and MCPM 2 comparisons were normally distributed.

These findings are contradictory to the findings of Ayaya (1996) and Zaaiman (1998) who found statistically significant positive relationships between academic performance and the completion of a bridging course. As less than 40 percent of the students in this sample on first and second-year level did the bridging course, it is recommended that more research is conducted to determine the impact that bridging course completion could have on academic performance in future years.

The last category where the differences between means were calculated was for the students who took commercial subjects at school (called Group 1 in table 5.17) versus the students who did not take commercial subjects at school (called Group 2 in table 5.17). The reference to commercial subjects used in this research project, represents students who did any one, two or all three of the following commercial subjects in matric:

- Business Economics
- Accountancy
- Mercantile law
- Economics

The aim was to determine whether students who do commercial subjects at school perform better than those who did not do commercial subjects at school.

TABLE 5.17 DIFFERENCE IN MCPM SCORES FOR STUDENTS WITH OR WITHOUT COMMERCIAL SUBJECTS AT SCHOOL, OVER A THREE-YEAR PERIOD

Academic year	F	Pr > F	d	N	Group 1		Group 2		
					Mean	Std dev	N	Mean	Std dev
First year	2.03	0.16	0.37 *	50	46.96	6.72	20	44.3	7.55
Second year	0.20	0.66	0.15 *	40	52.40	11.71	16	50.88	9.22
Third year	1.78	0.19	0.51 **	33	56.42	7.95	12	52.83	6.18

* d = 0.2 small effect

** d = 0.5 medium effect

*** d = 0.8 large effect

A test for normality reflected a normal distribution of the data for the MCPM 1 comparisons, but not for the MCPM 2 and MCPM 3 data. Therefore, only the t-test scores of the first year data will be interpreted. Students who did commercial subjects at school (Group 1) had slightly higher mean scores than those who did not do commercial subjects at school (Group 2), with a small practical significance effect. The differences were not statistically significant.

5.5 REGRESSION ANALYSIS

Regression equations provide a mathematical summary of the relationship between variables used in a regression analysis, which enable us to make predictions regarding the relationship between the variables included in the regression analysis (Tredoux & Durrheim, 2002). Both a multiple and stepwise regression analysis were conducted. An analysis of variance (ANOVA) was used for the regression analysis. ANOVA examines the separate and/or combined effects of two or more predictors or independent variables on a criterion or independent variable (Diekhoff, 1996; Tredoux & Durrheim, 2002). The

effects of various models, including different variables, on the academic performance of students (MCPM's) over the three-year period were investigated.

The following regression analyses were conducted

- Different independent variables were entered separately in a regression analysis against the primary and secondary criteria.
- A stepwise regression analysis using a model including all variables of table 4.2 with the primary and secondary criteria.

As the PIB and matric English did not correlate statistically significantly with tertiary academic results they were not subjected to the regression analysis, neither were, the bridging course status and commercial subjects in matric status data, as they also did not reflect any statistically significant difference in tertiary academic performance in the *t*-tests calculations. All variables were however included in the stepwise regression analysis.

5.5.1 The different independent variables entered separately in a regression analysis against the primary and secondary criterion

The different independent variables were entered separately with the MCPM 1 as the dependant variable. The results are provided in table 5.18.

TABLE 5.18 REGRESSION RESULTS OF DIFFERENT INDEPENDENT VARIABLES WITH MCPM 1 AS DEPENDENT VARIABLE

Independent variables	R²	Adjusted R-Square value	Significance of the F-value
LPCAT pretest	0.10	0.08	0.04*
LPCAT post-test	0.09	0.07	0.05*
Gender	0.04	0.02	0.12
MSPT	0.05	0.04	0.05*

The results in table 5.18 indicate that the LPCAT pretest and the LPCAT post-test, respectively, explained eight and seven percent of the variance of MCPM 1 at a statistically significant level. In the case of the LPCAT post-test the actual p-value was $p = 0.0484$ and was rounded off to 0.05, which is therefore statistically significant. MSPT explained four percent of the variance of MCPM 1. The statistical significance is a borderline case where the actual p-value was 0.0519 which was rounded off to 0.05.

The independent variables were also entered against the MCPM 2 as dependent variable and the results will be reported in table 5.19.

TABLE 5.19 REGRESSION RESULTS OF DIFFERENT INDEPENDENT VARIABLES WITH MCPM 2 AS DEPENDENT VARIABLE

Independent variables	R²	Adjusted R-Square value	Significance of the F-value
LPCAT pretest	0.25	0.22	0.00**
LPCAT post-test	0.21	0.19	0.01**
Gender	0.10	0.08	0.02*
MSPT	0.02	0.00	0.34

The results in table 5.19 indicate that the LPCAT pretest and post-test explained 22 and 19 percent of the variance in the MCPM 2 respectively, both at a statistically significant level. There is a significant increase in the extent to which the LPCAT pretest and post-test explain the variance from MCPM 1 to MCPM 2.

Gender explained eight percent of the variance of MCPM 2 at a statistically significant level. Note that in table 5.15 it was only on second year level that there was a statistically significant difference in the tertiary academic performance of males and females. As it is unconstitutional to base selection decisions on gender this information is not suitable to include in a selection battery.

The independent variables were also entered against the last category under the primary criterion, namely the MCPM 3 as dependent variable.

TABLE 5.20 REGRESSION RESULTS OF DIFFERENT INDEPENDENT VARIABLES WITH MCPM 3 AS DEPENDENT VARIABLE

Independent variables	R²	Adjusted R-Square value	Significance of the F-value
LPCAT pretest	0.29	0.26	0.00**
LPCAT post-test	0.29	0.26	0.00**
Gender	0.03	0.01	0.28
MSPT	0.09	0.07	0.04*

The results in table 5.20 reflect that for MCPM 3, the LPCAT pretest and LPCAT post-test each explained 26 percent of the variance of the dependent variable (MCPM 3) at a statistically significant level. The extent to which MSPT explains the variance in MCPM 3 increased to seven percent at a statistically significant level.

The similar extent in which the LPCAT pretest and post test explain the variance from MCPM 1 to MCPM 3 can be as a result of the multi-colinearity of these variables.

A comparison between tables 5.18 to 5.20 indicates an increase in the percentage that the LPCAT pretest and post-test explain the variance in academic performance over the three years of study. In table 5.18 the LPCAT pretest explains eight percent and the LPCAT post-test explains seven percent of the variance in MCPM 1. The degree in which it explained the variance increased to 26 percent of the variance in MCPM 3 (table 5.20). The biggest increase is however from first to second year level.

The extent, to which MSPT explains the variance of the primary criteria, also increases from four percent in the first-year to seven percent on third-year level.

Finally, the independent variables were entered separately against the secondary criterion, namely, the percentage subjects passed over a three-year period.

TABLE 5.21 REGRESSION RESULTS OF DIFFERENT INDEPENDENT VARIABLES WITH THE PERCENTAGE OF SUBJECTS PASSED OVER THREE YEARS AS DEPENDENT VARIABLE

Independent variables	R²	Adjusted R-Square value	Significance of the F-value
LPCAT pretest	0.17	0.15	0.00**
LPCAT post-test	0.16	0.14	0.00**
Gender	0.07	0.06	0.03*
MSPT	0.06	0.04	0.05*

In table 5.21 the LPCAT pretest and LPCAT post-test explained the highest degree of variance of the secondary criterion, namely, percentage subjects passed over three years, at a statistically significant level. Gender and MSPT also explained six and four percent of the variance in the percentage subjects passed over three years at a statistically significant level. The actual p-value of the MSPT regression is $p = 0.0511$ which was rounded off to 0.05 and it is therefore a borderline case of statistical significance.

5.5.2 Stepwise regression

A stepwise regression analysis, using a model including all variables against the primary and secondary criteria, was performed. The results of the stepwise multiple regression with all the variables, called the complete model, and MCPM 1 to MCPM 3 as predictors, are summarised in tables 5.22 to 5.33.

The variables that were included as part of the complete model in the stepwise regression analysis were:

- Matric symbol point total (MSPT)
- Matric English score
- Commercial subjects in matric status
- LPCAT pretest score
- LPCAT post-test score
- PIB Reading Comprehension sub-test score
- PIB Calculation sub-test score
- PIB Mental Alertness sub-test score
- Bridging course status
- Gender

5.5.2.1 Stepwise regression analysis of the complete model against the primary criteria

The different independent variables, listed in section 5.5.2, were entered into a stepwise regression analysis to determine which independent variable (or variables) jointly explained a significant degree of variance of MCPM 1 as the dependant variable. See the results in tables 5.22 to 5.24.

Note that due to the multi-collinearity of variables the percentage of variance explained may differ depending on whether a variable is entered together with other variables in a stepwise procedure or on its own (see previous section tables 5. 18 to 5.21).

TABLE 5.22 SUMMARY OF THE STEPWISE REGRESSION OF THE COMPLETE MODEL AGAINST MCPM 1

Model	Partial R-Square	R²
1	0.08	0.08 (a)
(a) LPCAT post-test		

TABLE 5.23 ANOVA OF REGRESSION OF THE COMPLETE MODEL AGAINST MCPM 1

	Sum of Squares	Df	Mean square	F	Significance
Regression	153.191	1	153.19	3.18	0.08
Residual	1781.17	37	48.13		
Total	1934.36	38			

TABLE 5.24 STANDARDISED COEFFICIENTS FOR THE COMPLETE MODEL AGAINST MCPM 1

Coefficients	Beta	Significance
LPCAT post-test	0.30	0.08

Only one variable, namely LPCAT post-test, emerged from the stepwise regression analysis, but did not explain the variance of MCPM 1 at a statistically significant level.

The different independent variables, listed in section 5.5.2, were then entered into the stepwise regression analysis to determine which an independent variable (or variables) explained the highest degree of variance of MCPM 2 as the dependant variable. See the results in tables 5.25 to 5.27.

TABLE 5.25 SUMMARY OF THE STEPWISE REGRESSION OF THE COMPLETE MODEL AGAINST MCPM 2

Step	Partial R-Square	R²
1	0.16	0.16(a)
2	0.08	0.24 (b)

Predictors: (a) LPCAT pretest and (b) Gender

TABLE 5.26 ANOVA OF REGRESSION OF THE COMPLETE MODEL AGAINST MCPM 2

	Sum of Squares	Df	Mean square	F	Significance
Regression	385.24	2	192.62	3.95	0.03
Residual	1219.19	25	48.77		
Total	1604.43	27			

TABLE 5.27 STANDARDISED COEFFICIENTS FOR THE COMPLETE MODEL AGAINST MCPM 2

Coefficients	Beta	Significance
LPCAT pretest	0.35	0.11
Gender	-4.63	0.11

In the stepwise regression analysis, which included all variables listed in section 5.5.2, the LPCAT pretest showed up from the stepwise regression analysis, explaining 16 percent of the variance in MCPM 2. In step two, gender appeared as a second variable increasing the explanation of variance in MCPM 2 to 24 percent. Although the two variables were not statistically significant on their own, the model, including the two variables, was statistically significant at $p = 0.03$.

Due to the multi-collinearity of the LPCAT pretest and post-test, already discussed in section 5.3.1, the individual independent variables do not have statistical significant values on its own, but, together, the variables reach statistically significant values.

As indicated in chapter three, it will be unconstitutional to consider gender as a potential element to include in a selection battery and therefore these results will not be considered for use in a selection model.

Finally the different independent variables, listed in section 5.5.2, were entered into the stepwise regression analysis to determine which independent variable (or variables jointly) explained a significant degree of the variance of the MCPM 3 as the dependant variable. See the results in tables 5.28 to 5.30.

TABLE 5.28 SUMMARY OF THE STEPWISE REGRESSION OF THE COMPLETE MODEL AGAINST MCPM 3

Step	Partial R Square	R²
1	0.35	0.35 (a)
Predictor: (a) LPCAT post-test		

TABLE 5.29 ANOVA OF REGRESSION OF THE COMPLETE MODEL 9ik AGAINST MCPM 3

	Sum of Squares	Df	Mean square	F	Significance
Regression	547.77	1	547.77	11.56	0.003
Residual	1042.06	22	47.37		
Total	1589.83	23			

TABLE 5.30 STANDARDISED COEFFICIENTS FOR THE COMPLETE MODEL AGAINST MCPM 3

Coefficients	Beta	Significance
LPCAT post-test	0.83	0.003

For MCPM 3 the LPCAT post-test was the only variable that was identified in the stepwise regression analysis, explaining 35 percent of the variance in MCPM 3, at a statistically significant level of $p < 0.005$.

5.5.2.2 Stepwise regression analysis of the complete model against the secondary criterion

The different independent variables were entered together into a stepwise regression analysis against the percentage subjects passed over three years, as the dependent variable. See the results in tables 5.31 to 5.33.

TABLE 5.31 SUMMARY OF THE STEPWISE REGRESSION OF THE COMPLETE MODEL AGAINST PERCENTAGE OF SUBJECTS PASSED OVER THREE YEARS

Model	Partial R Square	R²
1	0.17	0.17 (a)
2	0.08	0.24 (b)

Predictors: (a) LPCAT post-test, (b) Gender

TABLE 5.32 ANOVA OF REGRESSION OF THE COMPLETE MODEL AGAINST PERCENTAGE SUBJECTS PASSED OVER THREE YEARS

	Sum of Squares	Df	Mean square	F	Significance
Regression	10545	2	5272.60	5.68	0.007
Residual	33413	36	928.15		
Total	43959	38			

TABLE 5.33 STANDARDISED COEFFICIENTS FOR THE COMPLETE MODEL AGAINST PERCENTAGE OF SUBJECTS PASSED OVER THREE YEARS

Coefficients	Beta	Significance
LPCAT post-test	1.7	0.03
Gender	-19.38	0.07

In the first step of the stepwise regression analysis, the LPCAT post-test was identified, explaining 17 percent of the variance in the percentage of subjects passed over three years, statistically significant at $p < 0.05$.

A second variable appeared in step two of the stepwise analysis. In step two, gender increased the explanation in the variance of the percentage of subjects passed over the three-year period to 24 percent, but it was not statistically significant.

These positive and statistically significant explanations of the variance of MCPM 2, MCPM 3 and the percentage of subjects passed over three years indicate that the LPCAT is a valid predictor of academic performance for students in the National Diploma: Human Resources Management at the Nelspruit campus of Technikon Pretoria.

5.6 CHAPTER SUMMARY

This chapter contained information regarding correlation and regression calculations. The statistical results were reported and interpreted according to the literature review and compared with other similar research projects.

The results obtained in the statistical analysis of the data made it possible to address all the objectives of the study.

CHAPTER 6

INTERPRETATION OF RESEARCH RESULTS, LIMITATIONS AND RECOMMENDATIONS

6.1 INTRODUCTION

To be able to address the general objective of the study, namely, to ascertain the predictive validity of a selection battery for Technikon students, a literature review was conducted and reported regarding an overview of cognitive ability and its measurement, as well as the criteria to be used in student selection. This was discussed in chapters two and three. In chapter four the empirical part of the research process was outlined, focusing on the gathering of the data and the statistical techniques that were used to analyse the data. The statistical results regarding the predictive powers of various independent variables were reported in chapter five. In this chapter the statistical results will be interpreted in terms of the objectives of this study and the literature review.

At the end of this chapter reference will be made to the limitations of the study and suggestions will be made regarding additional research that is required to clarify problems.

6.2 INTERPRETATION OF THE RESEARCH RESULTS

The statistical results of chapter five will be interpreted according to the aims of the project as outlined in section 1.3. The first two aims of this research project concerned the literature review and the explanation of the empirical part of the research process as mentioned in section 6.1. The other aims will now be addressed.

One of the objectives of this study was to determine the extent to which matric results could be used in a selection battery to predict future tertiary academic performance (table 5.6, 5.18 to 5.20). There was a statistically significant correlation between MSPT (matric symbol point total) and MCPM 1. The magnitude of the correlation, in terms of effect size, was between small and medium. There was a statistically significant correlation between MSPT and MCPM 3, correlating at a slightly higher level than with MCPM 1. MSPT did not correlate with MCPM 2 at a statistically significant level; neither did MSPT reveal the variance in MCPM 2 at a statistically significant level. MSPT explained four percent of the variance of MCPM 1 and the degree to which it explained the variance in MSPT 3 increased to seven percent, both at a statistically significant level.

MSPT's correlation with the secondary criterion, namely the percentage of subjects passed over a three-year period, was a borderline case of statistical significance (table 5.7). MSPT explained four percent of the variance in the percentage of subjects passed over a three-year period at a statistically significant level (table 5.21).

Further investigations regarding the role of specific matric subjects, namely English and/or commercial subjects, on tertiary academic performance revealed that matric English correlated at a very low level which was not statistically significant with the primary and secondary criterion (tables 5.8 and 5.9). Neither did commercial subjects in matric show any statistically significant difference (table 5.17) in tertiary academic performance. This further supports a conclusion that matric results have limited predictive validity for tertiary academic performance.

These findings are in accordance with the findings of local research in this regard, such as the research of Huysamen (2000), Miller (1992), Shochet (1986), Skuy et al. (1996) and Zaaïman (1998), who also found that matric results did not accurately predict tertiary academic performance. It therefore appears as if matric results (MSPT) and/or individual, or grouped matric subjects are to a lesser degree valid and reliable in predicting tertiary academic performance and should not be used as a single selection

criterion until the schooling of all candidates in South Africa is on an equal base and prepares all students adequately for tertiary education.

Another objective of the study was to determine the predictive validity of two psychometric tests as predictors of future academic performance. A dynamic measure of learning potential, namely the LPCAT, and a static measure of intelligence, namely the PIB, were used.

Statistically significant Pearson correlation coefficients, with a medium to large effect, were found between the LPCAT pretest and post-test scores from first year to third year (table 5.10). The LPCAT pretest and post-test also correlated statistically significantly with the secondary criterion (table 5.11). This indicates that the LPCAT pretest and post-test correlate with tertiary academic performance of students in this study.

In the regression analysis, where all the independent variables were entered separately, the results also revealed that the LPCAT pretest and post-test respectively explained the highest variance of MCPM 1 to MCPM 3 (tables 5.18 to table 5.20) of all variables entered. The degree to which it explained the variance in the MCPM's increased from seven (post-test) and eight (pretest) percent on first-year level, increasing up to 26 percent (for both pretest and post-test) on third-year level. In the case of the secondary criterion (table 5.21), the LPCAT pretest and post-test respectively explained 15 and 14 percent of the variance in the number of subjects passed over the three-year period, at a statistically significant level.

In the stepwise regression analysis the LPCAT pretest and post-test did not explain the variance of MCPM 1 at a statistically significant level. On second year level, the LPCAT pretest, together with gender as variable, explained 24 percent of the variance in MCPM 2 at a statistically significant level (tables 5.25 to 5.27). On third year level only the LPCAT post-test emerged from the stepwise procedure and explained 35 percent of the variance in MCPM 3 at a statistically significant level (tables 5.28 to 5.30).

As indicated earlier the multi-colinearity of the LPCAT pretest and post-test could have affected the stepwise procedure, in so far that the individual variables might not have statistical significant values on its own, but together the variables explain the variances at statistical significant levels, like on second year level. The multi-colinearity can further influence the degree to which the LPCAT pretest and post-test explained the variance when it is entered separately, compared to when it is entered together with other variables, like in the stepwise regression procedure.

From these correlation and regression results it appears as if the LPCAT is a valid predictor of tertiary academic performance for the students in this sample. The LPCAT can therefore be regarded as a suitable measure to include in a selection battery for student selection purposes for the National Diploma: Human Resources Management at Technikon Pretoria Nelspruit campus.

The correlation values that were obtained from the correlation analysis, measuring the relationship between the PIB and tertiary academic performance, indicated low and/or negative correlations that were not statistically significant, between the PIB sub-tests and the three years' MCPM's (table 5.12). The correlation between the PIB sub-tests and the secondary criterion, namely, the percentage of subjects students pass over a three-year period were not statistically significant (table 5.13). In the stepwise regression analysis (tables 5.22 to 5.33) none of the PIB subtests explained the variance in the primary or secondary criteria at a statistically significant level. These findings are in accordance with the literature review reported in chapters two and three, which pointed out that static measures of cognitive ability are often not suitable to use as predictors for tertiary academic performance for students from an educationally disadvantaged background.

In comparing the predictive validity of the two psychometric tests, the results in this research project indicate that the learning potential measure, namely the LPCAT, is more effective and has a higher predictive validity in predicting future academic performance, than the static test used in this project, namely the PIB.

Another objective of this project was to determine whether the exposure to and completion of a bridging course had any effect on the performance of students. A *t*-test was conducted and revealed that the completion of a bridging course had no statistically significant impact on the tertiary academic performance (table 5.16) of students in any of the three years of study. As only 36 percent of the sample represented students who have completed a bridging course, it is recommended that this information should be used with caution and that research in this regard should rather be repeated on a bigger sample size to confirm or reject these findings.

The last objective of the study was to determine whether first year performance was related to subsequent years' performance. There were statistically significant correlations between the MCPM 1 and MCPM 2, as well as MCPM 2 and MCPM 3 which revealed that the students in the sample performed better once they had adjusted to the demands of tertiary education, confirming the late-blooming hypothesis (table 5.14). In terms of effect size the magnitude of these correlations were between medium and large. This was also revealed in the calculation of the differences in mean scores between MCPM of different semesters and years (table 5.3) where the effect size of the difference between second and first year performance was between medium and large, while the difference between third and second year performance were between small and medium. Huysamen (2000), who investigated whether results at university level were a better predictor for subsequent years' academic performance, also found that first year performance was a more valid predictor than matric marks in predicting future tertiary academic performance in the second and third years. This study reflects the same tendency.

6.3 LIMITATIONS

This research project is not without limitations.

Due to the relatively small sample size, limited conclusions are possible. There are, however, indications of both statistical and practical significance of results. The normality of the data was also tested and confirmed a normal distribution of data, except those of the bridging course and the commercial subjects at school data, as indicated earlier. The findings of this project are also consistent with those of other similar research projects based on larger sample sizes.

The results in this project are from one course within a single institution and can therefore not be generalised for other courses and/or other institutions.

All students came from the same racial group and therefore no cross-cultural information was available. Furthermore, all students in the sample matriculated from one educational authority and it was not possible to compare academic performance of students with different schooling backgrounds.

The number of subjects that a student registers for per year was not taken into account. This can have an effect on the correlations, as the course load of each student can affect the comparability of students' MCPM's.

The multi-colinearity of the LPCAT pretest and post-test could have influenced the regression results and one should consider only one in future regression analysis.

6.4 RECOMMENDATIONS

The above findings can have possible implications on student selection at Technikon Pretoria and other tertiary institutions facing similar dilemmas in student selection of previously disadvantaged students. As the Nelspruit satellite campus follows the same selection principles as the main campus, this research project will have to be repeated on

a larger sample of students registered for the National Diploma: Human Resources Management and other courses at the Pretoria campus and satellite campuses, to confirm similar findings before changes to the current selection process can be considered. This type of research should be undertaken continuously in all courses.

A larger sample size, as recommended above, should further include individuals from various racial groups who matriculate from various educational authorities to enable researchers to make more comparisons.

The number of subjects, which each student registers for per year, should be taken into account and a weighted mean should rather be calculated to prevent any influence on the correlation values.

In compiling a selection battery that reflects high predictive validity, it is recommended that other non-cognitive factors should also be considered, such as locus of control and study habits. It might also be worthwhile to include a test to assess the language proficiency of students.

This study, or similar research, should be repeated in other study areas with bigger sample sizes to determine the degree to which this information can contribute towards compiling a selection battery with high predictive validity.

6.5 CONCLUSION

The purpose of this study was to find valid predictors of academic performance that could be used in a selection battery for Technikon students. There was an indication that dynamic measures of cognitive ability have higher predictive validity than static psychometric tests and the other independent variables used in this project.

This study revealed that the practice of using matric results and static measures of cognitive ability, as part of a selection battery, will have to be reconsidered. Once the education system in South Africa is equal for all scholars, a selection criterion such as matric results can be used, but until that is attained, alternative selection batteries should be considered and investigated on a regular basis.

REFERENCE LIST

- Agar, D. (1990). Non-traditional students : Perceptions of problems which influence academic success. *Higher Education, 19*, 435-454.
- Allen, G.L., Woodward, E.A., Jones, R.J., Terranova, M. & Morgan, B.B. (1990). Measures of learning rate and scholastic aptitude as predictors of performance in training-oriented academic courses. *The Journal of General Psychology, 117* (3), 277-293.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillian.
- Anastasi, A. & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Atlink van den Berg, W.M.M. (1987). The evaluation of selection tests for upgrading programmes in Botswana and Swaziland. *International Journal of Educational Development, 7* (1), 1-12.
- Ayaya, O.O. (1996). The prediction of academic performance in the first year : A case study at the National University of Lesotho. *South African Journal of Higher Education, 10* (2), 101-113.
- Babad, E. & Budoff, M. (1974). Sensitivity and validity of learning potential measurement in three levels of ability. *Journal of Educational Psychology, 66*, 439-447.
- Baron, R.A. (1994). *Psychology* (3rd ed.). London: Prentice Hall.
- Berry, J.W. (1980). Cultural universality of any human intelligence remains an open question. *Behavioral & Brain Sciences, 3*, 584-585.

- Biesheuvel, S. (1943). *African intelligence*. Johannesburg: South African Institute of Race Relations.
- Biesheuvel, S. (1973). Adaptability : Its measurement and determinants. In L.J. Cronbach & P.J.D. Drenth (Eds.), *Mental tests and cultural adaptation* (pp.47-61). The Hague: Mouton.
- Boeyens, J.C.A. (1989a). Learning potential : A theoretical perspective. *Report PERS*, 432. Pretoria: Human Sciences Research Council. Pretoria.
- Boeyens, J.C.A. (1989b). Learning potential : An empirical investigation. *Report PERS*, 435. Pretoria: Human Sciences Research Council. Pretoria.
- Bokhorst, F.D., Foster, D.H. & Lea, S.J. (1992). Factors affecting academic performance in first-year psychology at the University of Cape Town. *South African Journal of Higher Education*, 6 (1), 59-65.
- Brody, N. (1992). *Intelligence* (2nd ed.) London: Academic Press.
- Brown, N.W. (1994). Cognitive, interest and personality variables predicting first-semester GPA. *Psychological Reports*, 74, 605-606.
- Brown, A.I. & Campione, J.C. (1986). Academic intelligence and learning ability. In R.J. Sternberg & D.K. Detterman (Eds.), *What is intelligence?* (pp. 39-44). Norwood: Ablex Publishing Corporation.
- Brown, A.L. & French, L.A. (1979). The zone of potential development : Implications for intelligence testing in the year 2000. *Intelligence*, 3, 255-273.

- Büchel, F.P. & Scharnhorst, U. (1993). The learning potential assessment device (LPAD): Discussion of theoretical and methodological problems. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment : Theoretical, methodological and practical issues* (pp. 83-111). Amsterdam: Swets & Zeitlinger.
- Budoff, M. (1974). *Learning potential and educability among the mentally retarded*. Final project No. 312312. Cambridge Mass.: Research Institute for educational problems, Cambridge, Mental Health Association.
- Budoff, M. (1987a). Measures for assessing learning potential. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (pp. 173-196). New York: Guilford Press.
- Budoff, M. (1987b). The validity of learning potential assessment. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (pp. 52-64). New York: Guilford Press.
- Campione, J.C. & Brown, A.L. (1987). Linking dynamic assessment with school achievement. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (pp. 82-139). New York: Guilford Press.
- Claassen, N.C.W. (1990a). *Die meting van intelligensie in verskillende groepe met die Algemene Skolastiese Aanlegtoets (ASAT)*. Pretoria: RGN.
- Claassen, N.C.W. (1990b). *Measuring intelligence with the General Scholastic Aptitude Test (GSAT)*. Paper presented at the 8th National Congress of the Psychological Association of South Africa. Pretoria: HSRC, 3–5 October 1990.

- Claassen, N.C.W. (1995). *Cross cultural assessment in the human sciences*. Paper delivered at a work session on the meaningful use of psychological and educational tests. HSRC, 24 October 1995.
- Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.
- Cole, M. & Scribner, S. (1978). The development of higher psychological processes. In M. Cole, V. John-Steiner, S. Scribner & E. Souberman (Eds.), *L.S. Vygotsky, Mind in society*. London: Harvard University Press.
- Constitution of the Republic of South Africa, No. 108 (1996). Government Gazette, 378 (17678).
- Coxhead, P. & Gupta, R.G. (1988). Construction of a test battery to measure learning potential. In R.G. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency : Recent developments in assessment* (pp. 118-137). Basingstoke: Macmillian.
- Cronbach, L.J. (1973). Judging how well a test measures. New concepts, new analysis. In L.J. Cronbach & P.J.D. Drenth (Eds.), *Mental tests and cultural adaptation* (pp. 413-425). The Hague: Mouton.
- Cronbach, L.J. (1990). *Essentials of psychometric testing* (5th ed.). New York: Harper Collins.
- Daniel, M.H. (1997). Intelligence testing : Status and trends. *American Psychologist*, 52 (10), 1038-1045.

- Das, J.P. (1987). Foreword. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (p. iv-xii). New York: Guilford Press.
- Dawes, P., Yeld, N. & Smith, M.J. (1999). Access, selection and admission to higher education : Maximising the use of the school leaving examination. *South African Journal of Higher Education*, 13 (3), 97-104.
- De Beer, M. (2000a). *Learning Potential Computerised Adaptive Test (LPCAT) : Technical Manual*. Pretoria: Production Printers (UNISA).
- De Beer, M. (2000b). *Learning Potential Computerised Adaptive test (LPCAT) : User's Manual*. Pretoria: Production Printers (UNISA).
- De Beer, M. (2000c). *The construction and evaluation of a dynamic computerised adaptive test for the measurement of learning potential*. Unpublished D.Litt. et Phil dissertation. University of South Africa, Pretoria.
- De Beer, M. (2002). *Utility of Learning Potential Computerised Adaptive Test (LPCAT) scores in predicting academic performance of bridging students : A comparison with other predictors*. Paper presented at the 5th Industrial Psychology Conference, South Africa, CSIR, 13-14 June 2002.
- De Beer, M. (2003). *A comparison of learning potential results at various educational levels*. Paper presented at the 6th Annual Society for Industrial and Organisational Psychology of South Africa (SIOPSA) conference, 25-27 June 2003.
- De Jager, A.C., Van Lingen, J.M. & Watson, A.S.R. (1997). *Increasing access for Black students through a special admission procedure : The UPE experience*. Paper presented at the 17th Annual Conference of the Society of Student Counseling in Southern Africa, Port Elizabeth.

- Department of Education (1997). *Education White Paper 3*. A programme for the Transformation of higher education. *General Notice 1196 in the Government Gazette 18207 of 1997-08-15*. Pretoria.
- Diekhoff, G.M. (1996). *Basic statistics for the social and behavioral sciences*. New Jersey: Prentice Hall.
- Embretson, S.E. (1987). Towards development to a psychometric approach. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (pp. 141-170). New York: Guilford Press.
- Employment Equity Act, No 55 (1998). *Government Gazette*, (19370).
- Erasmus, P.F. (1997). *Standardisation, validity and reliability of assessment tools in a fast changing culturally diverse society*. Krugersdorp: RNATA.
- Erasmus, P.F. & Minnaar, G.G. (1995). *Manual-Potential Index Battery and The Advanced Potential Index Battery*. Unpublished manuscript.
- Everette, J.E. (1991). Tertiary entrance predictors of first year university performance. *Australian Journal of Education*, 35(1), 24-40.
- Eysenck, H.J. (1971). *The IQ argument : Race, Intelligence and education*. New York: The Library Press.
- Eysenck, H.J. (1988). The concept of 'intelligence' : Useful or Useless. *Intelligence*, 12, 1-16.
- Ferguson, G.A. (1954). On learning and human ability. *Canadian Journal of Psychology*, 8 (2), 95-101.

- Ferguson, G.A. (1956). On transfer and abilities of man. *Canadian Journal of Psychology*, 10, 121-131.
- Feuerstein, R. (1973). Cognitive assessment of the socioculturally deprived child and adolescent. In L.J. Cronbach & P.J.D. Drenth (Eds.), *Mental tests and cultural adaptation* (pp. 265-275). The Hague: Mouton.
- Feuerstein, R. (1980). *Instrumental enrichment : An investigation program for cognitive modifiability*. Baltimore: University Park Press.
- Feuerstein, R., Feuerstein, R. & Gross, S. (1997). The learning potential assessment device. In D.P. Flanagan, J.L. Genshaft & P.L. Harrison (Eds.), *Contemporary intellectual assessment : Theories, tests, and issues* (pp. 297-313). New York: Guilford Press.
- Feuerstein, R., Rand, Y. & Hoffman, M.B. (1979). *The dynamic testing of retarded performance : The learning potential testing devise : Theory, instruments, and techniques*. Baltimore: University Park Press.
- Feuerstein, R., Rand, Y., Hoffman, M.B. & Miller, R. (1980). *Instrumental Enrichment : An intervention programme for cognitive modifiability*. Baltimore: University Park Press.
- Feuerstein, R., Rand, Y., Jensen, M.R., Kaniel, S. & Tzuriel, D. (1987). Prerequisites for assessment of learning potential : The LPAD model. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluate learning potential* (pp. 35-51). New York: Guilford Press.
- Feuerstein, R., Rand, J. & Rynders, J.E. (1988). *Don't accept me as I am : Helping "retarded" people to excel*. New York: Plenum Press.

- Fourie, C.M. (1991). Keuring van eerstejaar-universiteitstudiante. *South African Journal of Education*, 11 (4), 188-194.
- Foxcroft, C.D. (1997). Psychological testing in South Africa : Perspectives regarding ethical and fair practices. *European Journal of Psychological Assessment*, 13 (3), 229-235.
- Gardner, H. (1993). *Frames of mind. The theory of multiple intelligences* (10th ed.). New York: Basic Books.
- Gardner, J.B. (1989). Black students, an endangered species. *South African Journal of Higher Education*, 3 (2), 165-167.
- Gould, S.J. (1981). *The mismeasure of man*. New York: Norton.
- Gredler, G.R. (1988). Assessment of learning ability : A current appraisal. In R.M. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency : Recent developments in assessment* (pp. 39-63). Basingstoke: Macmillian.
- Gregory, R. (1986). Seeing Intelligence. In R.J. Sternberg & D.K. Detterman (Eds.), *What is intelligence?* (pp. 21-39). Norwood: Ablex Publishing Corporation.
- Grigorenko, E.L. & Sternberg, R. (1998). Dynamic testing. *Psychological Bulletin*, 124 (1), 75-111.
- Gupta, R.G. & Coxhead, P. (1988). Why assess learning potential? In R.J. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency : Recent developments in assessment* (pp. 1-19). Basingstoke: Macmillian.
- Guthke, J. (1992). Learning tests : The concept, main research findings, problems and trends. *Learning and individual differences*, 4 (2), 137-151.

- Guthke, J. (1993). Developments in learning potential assessment. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment : Theoretical, methodological and practical issues* (pp. 43-67). Amsterdam: Swets & Zeitlinger.
- Guthke, J. & Stein, H. (1996). Are learning tests the better version of intelligence tests? *European Journal of Psychological Assessment*, 12 (1), 1-13.
- Hamers, J.H.M. & Resing, W.C.M. (1993). Learning potential assessment introduction. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment : Theoretical, methodological and practical issues* (pp. 23-41). Amsterdam: Swets & Zeitlinger.
- Hamers, J.H.M. & Sijtsma, K. (1993). Learning potential assessment epilogue. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 365-372). Amsterdam: Swets & Zeitlinger.
- Heffner, C.L. (2004). Research methods for education, Psychology and Social Sciences [Online] Available from: <http://allpsych.com/researchmethods/validityreliability.html> [Accessed on : 24/05/2004].
- Hegarty, S. (1988). Learning ability and psychometric practice. In R.M. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency : Recent developments in assessment* (pp. 22-38). Basingstoke: Macmillan.
- Higher Education Act, No 101 (1997). *Government Gazette* (18515).

- Hilgard, E.R. (1989). The early years of intelligence measurement. In R.L. Linn (Ed.), *Intelligence measurement, theory, and public policy : Proceedings of a symposium in honor of Lloyd G. Humphreys* (pp. 7-28). Urbana: University of Illinois Press.
- Hilliard, A.G. (1990). Back to Binet : The case against the use of IQ tests in the schools. In T. Haute (Ed.), *Contemporary Education*, 61 (4), 184-189.
- Horn, J. (1989). Models of intelligence. In R.L. Linn (Ed.), *Intelligence : Measurement, theory, and public policy : Proceedings of a symposium in honor of Lloyd G. Humphreys* (pp. 29-37). Urbana: University of Illinois Press.
- Hoy, C. & Gregg, N. (1993). *Assessment : The special educator's role*. California: Brooks/Cole.
- Huysamen, G.K. (1996a). Fair and unbiased admission procedures for South African institutions of higher education. *South African Journal of Higher Education*, 10 (2), 199-207.
- Huysamen, G.K. (1996b). Matriculation marks and aptitude test scores as predictors of tertiary educational success in South Africa. *South African Journal of Higher Education*, 10 (1), 199-207.
- Huysamen, G.K. (1999). Psychometric explanations for the poor predictability of the tertiary-academic performance of educationally disadvantaged students. *South African Journal of Higher Education*, 13 (1), 132-138.
- Huysamen, G.K. (2000). The differential validity of matriculation and university performance as predictors of post-first-year performance. *South African Journal of Higher Education*, 15 (2), 146-151.

- Huysamen, G.K. (2001). Die verband tussen matriekprestasie en eerstejaarprestasie vir opeenvolgende innames aan dieselfde universiteit. *South African Journal of Higher Education*, 15 (3), 142-149.
- Huysamen, G.K. (2002). Die prestasievoorsigte in verskillende universitêre vakrigtings. *South African Journal of Higher Education*, 16 (3), 139-147.
- Huysamen, G.K. & Raubenheimer, J.E. (1999). Demographic-group differences in the prediction of tertiary-academic performance. *South African Journal of Higher Education*, 13 (1), 171-177.
- Jackson, I.M. & Young, D.A. (1988). Student selection using a model which could predict success in first-year biological studies at university. *South African Journal of Education*, 8 (3), 170-175.
- Jacobs, G.J. (1987). 'n Voorgestelde metode by die keuring van universiteitstudente. *Bulletin vir Dosente*, 19 (2), 25-40.
- Jawitz, J. (1995). Performance in first- and second-year engineering at UCT. *South African Journal of Higher Education*, 9 (1), 101-108.
- Jensen, A.R. (1969). Intelligence, learning ability and socioeconomic status. *The Journal of Special Education*, 3 (1), 23-33.
- Jensen, A.R. (1980). *Bias in mental testing*. London: Methuen.
- Jensen, M.R. & Feuerstein, R. (1987). Historical perspectives. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (pp. 379-441). New York: Guilford Press.

- Jensen, M.R., Feuerstein, R., Rand, R., Kaniel, S. & Tzuriel, D. (1988). Cultural differences and cultural deprivation : A theoretical framework for differential intervention. In R.M. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency : Recent developments in assessment* (pp. 64-88). Basingstoke: Macmillan.
- Jonker, E. (1994). *Die voorspelling van akademiese prestasie by gedepriveerde studente*. Unpublished M.SocSc thesis. Bloemfontein: University of the Orange Free State.
- Kirk, R.E. (1996). Practical significance : A concept whose time has come. *Educational and Psychological Measurement*, 56 (5), 746-759.
- Kotze, H.N. (1994). *Keuringsmodelle vir universiteitstudierigtings : 'n Psigometriese ondersoek*. Unpublished Ph.D. dissertation. Potchefstroom University for Christian Higher Education.
- Kotze, H.N., Van der Merwe, D. & Nel, A. (1996). *Culture fair selection procedures : The case of psychometrics*. Paper presented at the 16th Annual Conference of the Society of Student Counseling in Southern Africa, PREMOS. Pretoria.
- Kriel, H. (2001). *The validation of a potential assessment battery for engineering technology students*. Unpublished M.A. thesis. University of Pretoria.
- Labour Relations Act, No 66 (1995). *Government Gazette*, 366 (16861).
- Lidz, C.S. (1987). Historical perspectives. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (pp. 3-30). New York: Guilford Press.
- Louw, A.D. (1992). *Selection of prospective students for tertiary education : A creative response to turbulence and diversity of the nineties*. Paper presented at the 12th

- Annual Conference of the Society for Student Counseling in Southern Africa, Cape Town.
- Louw, A.D. (1993). Keuring van voornemende studente. *South African Journal of Higher Education*, 7 (3), 156-161.
- Louw, A.D. (1996). The use of confidential reports for the selection of prospective students for the Cape Technikon. *South African Journal of Education*, 16 (4), 194-200.
- Malekele, M.D. (1994). *Prediction of academic success of students in an academic development programme*. Unpublished M.A. thesis. Bloemfontein: University of the Orange Free State.
- Mattarazzo, J.D. (1990). Psychological assessment versus psychological testing. Validation from Binet to the school, clinic, and courtroom. *The American Psychologist*, 45 (9), 999-1017.
- Maxwell, G.S. (1996). Issues in selection for vocational education : Some lessons from the Queensland experience. *Assessment in Education*, 3 (1), 29-35.
- Miller, R. (1992). Double, double, toil and trouble : The problem of student selection. *South African Journal of Higher Education*, 6 (1), 98-104.
- Minick, N. (1987). Implications of Vygotsky's theories for dynamic assessment. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (pp. 116-139). New York: Guilford Press.
- Mji, A. (2002). What influences students to university education? Insights from the horse's mouth. *South African Journal of Higher Education*, 16 (2), 166-172.

- Mouton, J. (1998). *Understanding social research* (2nd ed.). Pretoria: Van Schaik.
- Murphy, R. (2002). *A review of South African research in the field of dynamic assessment*. Unpublished M.A. thesis. University of Pretoria.
- Nunns, C. & Ortlepp, K. (1994). Exploring predictors of academic success in psychology I and WITS university as an important component of fair student selection. *South African Journal of Psychology*, 24 (4), 201-207.
- Nzimande, B. (1995). *To test or not to test?* Paper presented at the Congress on Psychometrics. CSIR, Pretoria, 5–6 June 1995.
- Owen, K. (1998). *The role of psychological tests in education in South Africa. Issues, controversies and benefits*. Pretoria: HSRC.
- Papalia, D.E. & Olds, S.W. (1988). *Psychology (International student ed.)*. New York: McGraw-Hill.
- Poortinga, Y.H. (1995). Cultural bias in assessment : Historical and thematic issues. *European Journal of Psychological Assessment*, 11 (3), 140-146.
- Retief, A. (1988). *Method and theory in cross-cultural psychological assessment*. Pretoria: HSRC.
- Roe, R.A. (1989). Designing selection procedures. In P. Herriot (Ed.), *Assessment and selection in organizations : Methods and Practice for Recruitment and Appraisals*. New York: Wiley.
- Ruijsenaars, A.J.J.M., Castelijns, J.H.M. & Hamers, J.H.M. (1993). The validity of learning potential tests. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars

- (Eds.), *Learning potential assessment : Theoretical, methodological and practical issues*. Amsterdam: Swets & Zeitlinger.
- Russel, J.S., Persing, D.L., Dunn, J.A. & Rankin, R.J. (1990). Focusing on the selection process : A field study of the measures that influence admission decisions. *Educational and Psychological Measurement*, 50, 901-913.
- Rutherford, M. & Watson, P. (1990). Selection of students for science courses. *South African Journal of Education*, 10 (4), 353-359.
- Samkin, J.G. (1996). Should matriculation results represent the sole admission criterion for first-year accounting programme – Preliminary evidence from the University of Durban- Westville. *South African Journal of Education*, 16 (2), 117-122.
- Schaap, P. (1997). Guidelines for evaluating and validating assessment tools. In P.F. Erasmus (Ed.), *Unfair labour practise : The worker and Assessment*. Krugersdorp: RNATA.
- Seligman, L. (1994). *Developmental career counseling and assessment* (2nd ed.). California: Sage.
- Sewell, T.E. (1987). Dynamic assessment as a nondiscriminatory procedure. In C.S. Lidz (Ed.), *Dynamic assessment : An international approach to evaluating learning potential* (pp. 426-443). New York: Guilford Press.
- Shochet, I.M. (1986). *Manifest and potential performance in advantaged and disadvantaged students*. Unpublished Ph.D. dissertation. University of the Witwatersrand, Johannesburg.
- Shochet, I.M. (1992). A dynamic assessment for undergraduate admission: the inverse relationship between modifiability and predictability. In H.C. Haywood & D. Truziel (Eds.), *Interactive assessment* (pp. 332-355). New York: Springer-Verlag.

- Shochet, I.M. (1994). The moderator effect of cognitive modifiability on traditional undergraduate admissions test for disadvantaged black students in South Africa. *South African Journal of Psychology*, 24 (4), 208-215.
- Sijtsma, K. (1993). Psychometric issues in learning potential assessment. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars (Eds.), *Learning potential assessment : Theoretical, methodological and practical issues* (pp. 175-185). Amsterdam: Swets & Zeitlinger.
- Skuy, M., Zolezzi, S., Mentis, M., Fridjhon, P. & Cockcroft, K. (1996). Selection of advantaged and disadvantaged South African students for university admission. *South African Journal of Higher Education*, 19 (1), 10-118.
- Smith, E.H. (1992). *The outline for implementing or improving selection procedures for prospective technikon students, with specific reference to engineering courses*. Paper presented at the 12th Annual Conference of the Society for Student Counseling in Southern Africa, Cape Town.
- Spearman, C.E. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, 15, 201-293.
- Sternberg, R.J. (1985). *Beyond IQ. A triarchic theory of human intelligence*. Cambridge: Cambridge University Press.
- Sternberg, R.J. (1997). Intelligence and lifelong learning. *American Psychological Association*, 52 (10), 1134-1139.
- Stoker, D.J., Engelbrecht, C.S., Crowther, N.A.S., Du Toit, S. H. C. & Herbst, A. (1985). *Ondersoek na differensiële toelatingsvereisters tot tersiêre-onderwysinrigtings*. Pretoria: RGN.

- Sutton, A. (1988). L.S. Vygotskii. The cultural-historical theory, national minorities and the zone of next development. In R.M. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency : Recent developments in assessment* (pp. 89-116). Basingstoke: Macmillian.
- Taylor, T.R. (1994). A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework. *South African Journal of Psychology*, 21, 184-192.
- Taylor, T.R. & Boeyens, J.C.A. (1991). The comparability of the scores of black and whites on South African Personality Questionnaire : An exploratory study. *South African Journal of Psychology*, 21, 1-10.
- Tredoux, C. & Durrheim, K. (Eds.) (2002). *Numbers, hypotheses & conclusions : A course in statistics for the social sciences*. Landsdown, Cape Town: UCT Press.
- Tuckman, B.W. (1975). *Measuring educational outcomes : Fundamentals of testing*. New York: Harcourt Brace Jovanovich.
- Tzuriel, D., Samuels, M.T. & Feuerstein, R. (1988). Non-intellective factors in dynamic assessment. In R. G. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency : Recent developments in assessment* (pp. 141-158). Basingstoke: Macmillian.
- Uys, L.R. (1994). Selection criteria predictive of academic success in the first two years of integrated diploma in Nursing. *Curationis*, 16 (2), 1-4.
- Van Aswegen, M. (1997). *The standardisation of a learning potential battery for the selection of poorly qualified employees*. Unpublished M.Com. thesis. University of Pretoria.

- Van de Vijver, F.J.R. & Poortinga, Y.H. (1987). Explaining cross-cultural differences : Bias analysis and beyond. *Journal of Cross-cultural Psychology*, 18 (3), 259-282.
- Van de Vijver, F.J.R. & Poortinga, Y.H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13 (1), 29-37.
- Van de Vijver, F.J.R. & Tanzer, N.K. (1997). Bias and equivalence in cross-cultural assessment : An overview. *European Review of Applied Psychology*, 47 (4), 263-279.
- Van den Berg, A.R. (1989). Basic item response theory. *IPEER training series*, 2585. Pretoria : Human Sciences Research Council.
- Van Eeden, R., De Beer, M. & Coetzee, H.C. (2001). Cognitive ability, learning potential, and personality traits as predictors of academic achievement by engineering and other science and technology students. *South African Journal of Higher Education*, 15 (1), 171-179.
- Van Niekerk, H.A. (1991). *Evaluation of Feuerstein's instrumental enrichment programme for culturally different senior students*. Unpublished M.Ed. thesis. University of Pretoria.
- Van Rooyen, E. (2001). Die voorspelling van die akademiese prestasie van studente in 'n universiteitsoorbruggingsprogram. *South African Journal of Higher Education*, 15 (1), 180-188.
- Van Zyl, M.A., Terblanche, S. & Jacobs, G.J. (1992). Die profiel van 'n potensieel-suksesvolle eerstejaarstudent in maatskaplike werk : 'n Benadering tot keuring. *Social Work*, 28 (1), 33-43.

- Verster, J.M. (1983). Human cognition and intelligence : An integrated theoretical perspective. *Report PERS, 350*. Pretoria: Human Sciences Research Council.
- Verster, J.M. (1985). *Bias and fairness in psychological assessment : clarification of terms*. Unpublished document, HSRC workshop on bias and fairness in psychological testing, Pretoria.
- Verster, J.M. (1988). Cognitive competence in Africa and models of information processing : A research prospectus. In J.W. Berry, S.H. Irvine & E.B. Hunt (Eds.), *Indigenous cognition : Functioning in a cultural context* (pp. 127-156). NATO ASI Series D No. 41. Dordrecht: Nijhoff.
- Vygotsky, L.S. (1978). *Mind in society. The development of Higher Psychological Processes*. Cambridge: Harvard University Press.
- Wegner, T. (1993). *Applied statistics*. Cape Town: Juta.
- Weiten, W. (1994). *Psychology : Themes and Variations* (2nd ed.). California: Pacific Grove.
- Wilson, K.M. (1980). The performance of minority students beyond the freshmen year : Testing the “late bloomer” hypothesis in one state university setting. *Research in Higher Education, 13* (1), 23-47.
- Wilson, K.M. (1983). A review of research on the prediction of academic performance after the freshman year. *College Board Report, 83* (2). New York: College Board Publications.
- Zaaiman, H. (1998). *Selecting students for mathematics and science. The challenge facing higher education in South Africa*. Unpublished Ph.D. thesis. Amsterdam: Vrije Universiteit.

Zaaiman, H., Van der Flier, H. & Thijs, G.D. (1998). Selecting South African higher education students : critical issues and proposed solutions. *South African Journal of Higher Education*, 12 (3), 96-101.

Zeidner, M. (1988). Culture fairness in aptitude testing revisited : A cross-cultural parallel. *Professional Psychology: Research and Practice*, 19 (3), 257-262.

Zolezzi, S.A. (1992). *Alternative selection measures for university undergraduate admissions*. Unpublished M.Ed. dissertation. University of Witwatersrand, Johannesburg.

Zolezzi, S.A. (1995). *The effectiveness of dynamic assessment as alternative aptitude testing strategy*. Unpublished D.Phil. dissertation. University of South Africa, Pretoria.