THE VALIDATION OF A POTENTIAL ASSESSMENT BATTERY FOR ENGINEERING TECHNOLOGY STUDENTS

by

Helena Kriel

submitted in fulfilment of the requirements
for the degree of

Magister Artium
(Psychology)

in the

Faculty of Humanities

at the

University of Pretoria

Pretoria **June 2001**

Dedicated to my parents Maarten and Alta Olivier,
for relentlessly believing in me

and

to the loving memory of Mariana Bothma, for teaching me the art of being alive

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Summary

THE VALIDATION OF A POTENTIAL ASSESSMENT BATTERY FOR ENGINEERING TECHNOLOGY STUDENTS

by

Helena Kriel

Study leader: Prof RP de la Rey
Department: Psychology
Degree: Magister Artium (Psychology)

The selection of students for higher education has been a burning issue on the agenda of South African institutions of higher education for the past decade. Institutions for higher education are experiencing pressure from both their clients and the government to broaden access, but at the same time financial realities force these institutions to admit only those candidates with the potential to be successful in their chosen course of study.

The main aim of this study was the identification of variables which relate to academic success amongst Engineering Technology students at Technikon Pretoria, and to incorporate them into a selection battery which would be both valid and reliable.

A non-experimental, correlational design was selected, as this research technique is considered the best controlled and most accurate of all non-experimental designs. Since a quantitative technique was selected for data gathering, the necessity for a statistical method in the data analysing process was obvious.

The sample for this study consisted of a total of 732 Engineering Technology students at Technikon Pretoria. From these, 512 were Civil Engineering Technology students and the remaining 220 were Mechanical Engineering Technology students. These subjects were the total number of students from these two academic departments, enrolled from 1997 to 1999, of whom both psychometric and academic data were available. The sample consisted of 14.75% female and 85.25% male respondents and was representative of the cultural diversity of the Technikon campus.

The competencies indicated by academic staff involved with the training of Engineering Technology students at Technikon Pretoria were hypothesised to be indicative of a potentially successful student. After the identification of these predictor variables the assessment battery to be used in this study was compiled. This was then included in a comprehensive set of data regarding each applicant, together with the required school performance. A forward stepwise multiple regression analysis was performed on the data in order to establish the predictive validity of the assessment battery.

The expansion of the traditional selection procedure to include the potential assessment phase proved valuable, as the validity of all prediction models improved with the addition of the indices from the Potential Index Batteries. The prediction models were found to be unbiased against students from the previously disadvantaged school systems and can thus be said to be culture fair.

Key terms: Selection; assessment; assessment battery; potential; ability; competency; validity; reliability; regression model.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

OPSOMMING

DIE VALIDERING VAN 'N POTENSIAALBEPALINGSBATTERY VIR INGENIEURSTEGNOLOGIESTUDENTE

deur

Helena Kriel

Studieleier: Prof RP de la Rey
Departement: Sielkunde
Graad: Magister Artium (Sielkunde)

Die keuring van studente is vir die afgelope dekade 'n belangrike punt op die agenda van hoëronderwysinstellings in Suid-Afrika. Hoëronderwysinstansies ondervind druk van beide hul kliënte en die regering van die dag om toelating tot die instansies te verhoog, maar terselftertyd noop finansiële realiteite hierdie instansies om slegs studente toe te laat wat oor die potensiaal beskik om suksesvol in die kursus van hul keuse te wees.

Die hoofdoel van hierdie ondersoek was om die veranderlikes wat met die akademiese sukses van Ingenieurstegnologiestudente verband hou, te identifiseer en in 'n geldige en betroubare keuringsprogram op te neem.

Daar is gebruik gemaak van 'n nie-eksperimentele, korrelatiewe ontwerp, aangesien hierdie ontwerp beskou word as die mees akkurate van alle nie-eksperimentele ontwerpe. Daar 'n kwantitatiewe tegniek vir data-insamaling verkies is, was die belangrikheid van 'n statistiese tegniek vir die verwerking van die data voor die hand liggend.

Die steekproef vir die studie het bestaan uit 732 Ingenieurstegnologiestudente aan Technikon Pretoria. Hiervan was 512 Siviele Ingenieurstegnologiestudente en die

oorblywende 220 Meganiese Ingenieurstegnologiestudente. Hierdie proefpersone was die totale groep studente uit hierdie twee departemente, geregistreer vir die tydperk 1997 tot 1999, vir wie daar beide psigometriese data en akademiese resultate beskikbaar was. Die steekproef het bestaan uit 14.75% vroulike en 85.25% manlike respondente en was verteenwoordigend van die kulturele diversiteit van die Technikonkampus.

Die bevoegdhede wat deur akademiese personeel betrokke by die opleiding van Ingenieurstegnologiestudente aangedui is, is hipoteties gestel as aanduidend van 'n potensieel suksesvolle student. Nadat hierdie voorpellingsveranderlikes geïdentifiseer is, is die evalueringsbattery vir gebruik in hierdie studie saamgestel. Die psigometriese data is ingesluit in 'n omvattende databasis ten opsigte van elke kandidaat, tesame met die vereiste skoolprestasie. 'n Voorwaardse stapsgewyse regressie-ontleding is op die data uitgevoer ten einde die voorspellingsgeldigheid van die potensiaalbepalingsbattery vas te stel.

Die uitbreiding van die tradisionele keuringsprosedure deur die insluiting van die potensiaalbepalingsfase het waardevol geblyk te wees, aangesien die geldigheid van alle voorspellingsmodelle verhoog het met die insluiting van die indekse van die Potential Index Batteries. Die voorspellingsmodelle is bewys synde nie sydig te wees ten opsigte van studente uit voorheen benadeelde skoolstelsels nie en kan daarom as kultuurbillik beskou word.

Sleutelterme: Keuring; evaluering; evalueringsbattery; potensiaal; vermoë; bevoegdheid; geldigheid; betroubaarheid; regressiemodel.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Acknowledgments

My heartfelt gratitude to the following:

- **Prof Piet de la Rey**, my study leader, for his knowledgeable inputs and guidance.
- My husband **Coenie**, for his unfaltering support and encouragement, without which this dissertation would never have seen the light.
- My parents **Maarten and Alta Olivier**, for their lifelong interest in my academic career and their infallible belief in my abilities.
- My parents-in-law **Ryno and Heléne Kriel**, for their support.
- All the **Olivier and Kriel siblings**, for being interested.
- My grandmothers, **Pieta Coertze** and **Sannie Olivier**, for always taking pride in me.
- My colleagues at the Department of Student Counselling: **Elmarie, Elani, Francette, Ilzé, Letta, Lorika and Marlize**, for all they did to give me the opportunity to complete this study.
- My former manager and mentor, the late **Mariana Bothma**, for the initiative that lead to the undertaking of this study.
- **Ingrid Swanepoel**, for capable and meticulous revision of grammar and style.

- **The NRF for the financial support that made this study possible. The views expressed in this study are those of the author, and not necessarily that of the NRF.**

SOLE DEO GLORIA !

Chapter 1

Introduction

1.1 Introduction

The selection of students for higher education has been high on the agenda of South African institutions of higher education for the past decade. Louw (1992:1) states that the criteria used in the selection of prospective students for tertiary studies are being questioned more and more.

In addition, institutions for higher education have experienced an increase in applicant numbers since 1994, which soon led to demand exceeding available vacancies, and hence some sort of selection became essential. This situation is further complicated by the fact that there is limited public funding available for higher education, and by the marked differences in terms of their preparedness for higher education that exist amongst students from the previously differentiated school systems (Huysamen, 1996:199; Zaaiman, 1998:1).

Furthermore, institutions of higher education are being pressurised by government to increase demographic representation in their student populations, as well as to address the developmental needs of society better by rectifying the mismatch between the output of higher education and the needs of the modernising economy (Department of Education, 1997:8). In this respect, specific mention is made of the shortage of highly trained graduates in the fields of engineering, science, technology and commerce.

In addition to the above, there is the problem of extremely high attrition rates experienced by institutions of higher education (Stumph, 1997:1). At Technikon Pretoria an average attrition rate of 34% has been recorded for first-year students in the period 1996 to 1999. This does not include the 35% of first-year students that failed more than half of their subjects.

Apart from the above considerations, there is the emotional suffering of those students who, through no fault of their own and despite hard work, simply are incapable of academic success at tertiary level (Huysamen, 1996:200).

It thus became imperative for a more effective system for managing access to be developed. Technikon Pretoria decided to investigate and develop an alternative selection strategy by means of which access could be increased. It was deemed extremely important that, whatever strategy was decided upon it should be fair, unbiased and effective.

Traditionally, Technikon Pretoria, as most other institutions of higher education in South Africa, has based its admission of students on scholastic academic performance. A Senior Certificate (or equivalent) was set as the absolute minimum requirement for admission. In some instances additional requirements were set, such as a minimum mark for Mathematic and Physical Science in the case of engineering technology courses.

As engineering is stated by government to be one of the focus areas for increased participation, it was decided that this study would look into a valid and reliable process by which potentially successful students could be admitted to Technikon Pretoria's Faculty of Engineering.

1.2 Objectives of the study

The main aim of this study was to develop a potential assessment battery for the selection of engineering technology students at Technikon Pretoria, which would be both valid and reliable.

The second objective of this study was to deduce general selection guidelines for institutions of higher education from the research experiences and results.

1.3 Elucidation of key terms

In order to enhance the clear interpretation of the key terms used in this dissertation, those terms are explained below.

Selection

For the purpose of this dissertation the term "selection" will be used to refer to the process by which information regarding an applicant is gathered and evaluated in order to decide on his or her suitability for admission to higher education. Chapter 3 deals with this aspect in detail.

Assessment/selection battery

The set of subtests compiled for use in the determinating of an applicant's potential to be successful in the relevant academic programme.

Potential

In this study "potential" is used to refer to the assumption that the ability to complete a simple task, indicates the potential to complete a more complicated, but related task. This is in accordance with verbal communication with Dr Pieter Erasmus, developer of the psychometric instrument used in this study.

Ability

For the purpose of this dissertation, the term "ability" refers to the capability to complete a task that has been set successfully.

Competency

This study concurs with the definition of Armstrong (1996:195) that competency refers to those behavioural characteristics needed to attain the required levels of performance.

Engineering Technology student

A student enrolled for a course in either Mechanical Engineering or Civil Engineering Technology at Technikon Pretoria.

1.4 Arrangement of chapters

In conclusion, a short description of the arrangement of the chapters of this study is given.

Chapter 2 deals with the debate on the admission of students to higher education: problems faced by institutions for higher education, as well as policies and demands from government, are looked into. Chapter 3 addresses the concept of selection in detail and Chapter 4 examines psychometric testing, and briefly looks at the related political, legal and scientific issues. Chapter 5 deals with the development of criterion variables. Predictors of academic success, as determined by previous studies, are explored in Chapter 6. Chapter 7 describes the Potential Index Batteries, the psychometric instrument used in this study, and Chapter 8 deals with the research methodology followed in the execution of the investigation. The sample used in the study is described in Chapter 9 and the statistic techniques used in the analysis of the data are described in Chapter 10. The determination of the predictor variables are dealt with in Chapter 11. The results of the data analysis process are given and discussed in Chapter 12. Finally, Chapter 13 deals with conclusions and recommendations. Suggestions for future studies are also made in this chapter.

Chapter 2

The admission of students to higher education

2.1 Introduction

Section 32(b) of the Constitution of the Republic of South Africa, 1996 (Act No. 108 of 1996), states that every person has the right "to further education, which the state, through reasonable measures, must make progressively available and accessible."

In addition, the Reconstruction and Development Plan (1994:66) of the African National Congress, states that the higher education system represents a major resource for national development and contributes to the worldwide advancement of knowledge. The central role higher education plays in the social, cultural and economic development in modern societies is emphasised by the White Paper on Higher Education (Department of Education, 1997:7).

The matter of access to higher education played a major role in the events that led to this study, and therefore, attention will be given in this chapter to admission requirements for higher education as well as to the participation rate in higher education. As some of the issues addressed by the White Paper on Higher Education impact directly on this study, it will be discussed briefly.

2.2 Admission requirements for higher education

Present admission requirements for higher education consist of a Senior Certificate as well as matric exemption, in the case of universities. All higher education institutions regulate additional entrance requirements autonomously. In many instances, degree and diploma programmes involving mathematics and science have such extended requirements, for example Engineering, Chemistry, Biotechnology and Food Technology. In some professional fields, such as Physiotherapy, admission is determined by the relevant professional body and limitations are placed on the number of students admitted for study.

Amongst others the minimum requirements set by Technikon Pretoria for admission to the engineering technology courses will be looked into by this study.

2.3 Participation in higher education

In 1993, the overall higher education participation rate in all post-grade 12 educational programmes in public and private education institutions was estimated at 20%. The participation rates of Whites at this time was "slightly less than 70%, while that for Indians was about 40%, for Coloureds 13% and for Africans about 12%" (Stumph, 1997:1). The discrepancies between the distribution of participation can clearly be seen, and Moody (1994:7) states that "access is the first step in the cycle of increasing black participation in higher education."

2.3.1 Increasing access by means of alternative/special admission methods

De Jager et al (1997:2) state that the term *special admission* is used internationally for the processes used in combination with regular admission procedures, in order to accommodate applicants who would not qualify for admission, should regular criteria be applied. Special admission procedures are usually aimed at students whose matric results are below the standard required for admission at institutions for higher education , but who have the potential to benefit from higher education.

In the South African context the need for such an alternative admission process was created by the fragmented nature of the secondary school system of the past, as one of the implications of the inequalities in educational provision is the question of access (Kotze et al, 1996:39). This was caused by the fact that, traditionally, scholastic performance was used as major criterion to gain admission to higher education in South Africa.

An opposed opinion is offered by Huysamen (1997:65):

> "In several countries, vehement criticisms have been raised against the use of admission tests. In South Africa, the possibility of admission testing has acquired an additional dimension insofar as it is believed that it would be harmful to efforts to increase tertiary-educational access for previously excluded demographic groups."

At Technikon Pretoria, where this study was executed, a system was developed whereby a prospective student is evaluated in terms of his or her potential to be successful in a chosen course. The Technikon Pretoria Potential Assessment (TPPA), as the process is known, is seen as a way of levelling the playing fields; of giving all applicants an equal chance of access, irrespective of their past academic achievements. In this way Technikon Pretoria has taken the "first step in the cycle of increasing black participation in Higher Education" (Moody, 1994:1).

2.4 The White Paper on Higher Education

The need for the reconstruction of higher education in South Africa, which resulted from racial and gender inequalities and apartheid-based education, has been given priority by the ANC government, and the National Commission for Higher Education was established by President Mandela in February 1995.

The publication of the White Paper on Higher Education, which outlines a comprehensive set of initiatives for the transformation of higher education, was the result of the activities of this Commission.

The White Paper on Higher Education sets a programme for the transformation of higher education in South Africa. It focuses closely on the purposes and challenges of such higher education (Department of Education, 1997:7-14). These aspects will subsequently be discussed.

2.4.1 The purposes of higher education

According to the White Paper on Higher Education (Department of Education, 1997:7), the purposes of higher education are as follows:

- Higher education should meet the learning needs and aspirations of individuals through the development of their intellectual abilities and aptitudes throughout their lives. It is therefore an important vehicle for achieving equity in the distribution of opportunity and achievement among South African citizens.

- Higher education should address the development needs of society and provide the labour market with the ever changing high-level competencies and expertise necessary for the growth of a modern economy.

- Higher education should contribute to the socialisation of enlightened, responsible and constructively critical citizens.

- Higher education should contribute to the creation, sharing and evaluation of knowledge.

2.4.2 The needs of and the challenges facing higher education

The present system of higher education is limited in its ability to meet the moral, political, social and economic demands of the new South Africa (Department of Education 1997:8). According to the White Paper on Higher Education, the following deficiencies, *inter alia,* characterise the current higher education system:

- There is an inequitable distribution of access and opportunity for students and staff in terms of race, gender, class and geography and there are equally untenable disparities between historically white institutions and historically black institutions in terms of facilities and capacities.

- There is a chronic mismatch between the output of higher education and the needs of the modernising economy. There is a particular shortage of highly trained graduates in fields such as science, engineering, technology and commerce.

- Higher education has not adequately fulfilled its unmatched obligation to help lay the foundations of a critical civil society, with a culture of public debate and tolerance which accommodates differences and competing interests.

2.4.3 Requirements for the transformation of higher education

According to the White Paper on Higher Education, the following requirements will have to be met to bring about a successful transformation of the higher education system and its institutions (Department of Education, 1997:10):

- *Increased and broadened participation.* Successful policy should increase access for Blacks and women; disabled and mature students.

- *Responsiveness to societal interests and needs.* Institutions should meet the needs of an increasingly technology-orientated economy.

- *Cooperation and partnerships in governance.* The relationship between the higher institutions and the state, civil society and stakeholders as well as those between institutions themselves will have to be considered.

The emphasis placed on the redress of the inequitable distribution of access in terms of race and gender by the White Paper on Higher Education had a strong influence on the current study.

In the past, the admission of students on grounds of scholastic achievement only placed many prospective students in a disadvantaged position (Kotze et al, 1996:39) and led to the comment on "inequitable distribution of access" by the White Paper on Higher Education (Department of Education, 1997:8). The development of selection methods that will make higher education more accessible for different race and gender groups is therefore in accordance with the guidelines set for the transformation of higher education by the Department of Education.

In addition, the shortage of qualified graduates in engineering and technical fields addressed by the White Paper on Higher Education (Department of Education, 1997:8) also impacts on this study. The difficulties experienced by the Engineering Faculty at Technikon Pretoria in producing the required number of qualified technicians will be discussed in Chapter 6. Suffice it to say here that admitting the student with the best potential to be successful in the course of his or her choice, would help to rectify the mismatch between the output of higher education and the needs of the economy, as addressed by the White Paper.

2.4.4 Vision

The Ministry of Education envisages a system of higher education that will, *inter alia*, "promote equity of access and fair chances of success to all who are seeking to realise their potential through Higher Education, while eradicating all forms of unfair discrimination and advancing redress for past inequities." (Department of Education, 1997:11.)

2.4.5 Principles

According to the White Paper on Higher Education (Department of Education, 1997:11), the Ministry of Education regards eight principles as fundamental. However, only the four principles that are relevant to this study will be discussed here.

- **Equity and redress.**
  This principle implies fair opportunities for entering higher education programmes and succeeding in them. Such transformation would abolish all

existing forms of unfair differentiation.

- **Quality**
  The quality principle refers to the maintenance and application of academic and educational standards, "both in the sense of specific expectations and requirements that should be complied with, and in the sense of ideals of excellence that should be aimed at." (Department of Education, 1997:12.)

- **Institutional autonomy**
  The principle of institutional autonomy implies a high degree of self-regulation as far as, *inter alia,* student admission, curriculation, methods of teaching and assessment are concerned.

- **Public accountability**
  "The principle of public accountability implies that institutions are answerable for their actions and decisions not only to their own governing bodies and the institutional community, but also to the broader society." (Department of Education, 1997:13.)

The influence of the above principles on the study conducted here is clear: a selection battery which has predictive validity should give all a fair opportunity for entering higher education, and the success rate of the students admitted should be predictable. Furthermore, the quality of academic and educational standards would be determined by the potential of the student admitted. By using a battery such as the one suggested in this study, the institution would regulate its own student admissions and be publically accountable, as empirical research results regarding the predictive validity will be readily available.

2.5 Implications of legislation on current practices in higher education

According to Stumph (1997:5), the White Paper on Higher Education contains a "clear and unambiguous commitment to the promotion of equity of access and fair chances of success to all who are seeking to realize their potential through Higher Education."

Bodibe (1997:6) states that the management of access will, in future, imply moving away from exclusive access to access which guarantees equal opportunities to all groups previously disadvantaged.

In addition, according to the National Qualifications Framework, provision will have to be made for appropriate entry points for all prospective learners "in a way that facilitates progression" (Human Sciences Research Council, 1995:1). Recognition will also have to be given for all prior learning and other relevant competencies.

In the development of a potential assessment battery for prospective students evidence will have to be accumulated towards the principle of broadening access. Equal opportunities for access will have to be guaranteed. Furthermore, in the final selection, prior learning and industry related-experience that may enhance a student's chances to be successful, will have to be recognised.

2.5.1 Curriculum 2005

The implementation of Curriculum 2005 by the Department of Education will lead to the phasing out of the National Senior Certificate, which is currently the requirement for admission to higher education institutions. A pass in the Further Education Training

Certificate (FETC) will in future be the minimum statutory requirement for admission to higher education (Stumph, 1997:8). Institutions will however continue to have the right to determine entry requirements where they consider it to be appropriate over and above this statutory minimum. In affording them this right, the White Paper on Higher Education explicitly emphasises that such additional requirements should ensure that selection criteria are sensitive to the educational backgrounds of learners. Furthermore, such selection criteria should "incorporate the recognition of appropriate prior learning, including learning that took place in less structured and formalized environments than those typically encountered in our education system only" (Stumph, 1997:9).

2.6 Financial implications of student pass rates

As higher education institutions receive the bulk of their government subsidies for successful students, the "increasing financial burdens of Higher Education institutions likewise imply that institutions will no longer be able to bear the burden of unsatisfactory study progress by students" (Stumph, 1997:14). Pienaar (1991:9) agrees that higher education is cost-intensive and therefore it will become more and more important to select only students for higher education who have a chance of being successful.

According to Stumph (1997:2) the official average pass rate for students enrolled at South African technikons for three year diplomas was 15% for historically white institutions and 9% for historically black institutions in 1990. Official records for Technikon Pretoria show that 35% of the students registered as first year students in 1996 dropped out of their courses by the beginning of 1997. As was said earlier, this figure does not take into account the 34% of first-year students that failed more than half their subjects (Technikon Pretoria: Strategic Information and Planning, 1998).

2.7 Conclusion

The face of higher education in South Africa is set to change. The Constitution of the country gives all people the right to further education through it the State has committed itself to broadened access and availability.

Furthermore, the Department of Education has undertaken in its White Paper on Higher Education to address the mismatch between the "output of Higher Education and the needs of the modernising economy" (1997:8). In terms of the White Paper, this specifically refers to scientific, engineering and technology fields.

Hand in hand with the principle of the maintaining of standards goes the financial burden that unsuccessful students place on institutions for higher education. Higher education institutions could either drop their standards and in so doing increase their throughput statistics to increase their government subsidies, or maintain their standards, but ensure that the students they admit have the potential to be successful in their chosen courses.

The implementation of Curriculum 2005 is furthermore forcing higher education institutions to reconsider their present admission requirements and investigate alternative admission procedures.

From all the arguments above, it is clear that some innovative measures have to be developed to enable higher education institutions to select potentially successful students. A potential assessment programme with scientifically proven predictive validity is suggested as such a measure to ensure that those students with the best chance to pass would be admitted to higher education and that the institution concerned would benefit financially from the achievements of its students.

Chapter 3

Selection

3.1 Introduction

Kotze (1994:15) states that various aspects should be taken into consideration when a selection process for higher education is developed. In this chapter attention will be given to theoretical principles involved in designing a selection process and the application of those principles to a prediction model. Subsequently, the concept of fairness in selection will be discussed. In conclusion, a close look will be taken at performance prediction models.

3.2 Defining the concept

Blake (1983:237) defines selection as "the process of choosing, from those available, the person or people who best meet the requirements of a position or positions vacant within an organization."

Leap and Crino (1993:237) continue by stating that the purpose of selection is to discriminate fairly between applicants. "A firm must be able to separate applicants who will perform well as employees from applicants who will not" (Leap and Crino, 1993:237).

In the context of this study, "selection" will refer to the process of choosing the potentially successful student rather than the potentially successful employee.

3.3 Principles underlying the development of a selection procedure

During the past decades that make up the history of personnel selection, little attention has been given to the construction of integral selection procedures to be used in practical conditions (Roe & Greuter, 1991:190). In order to address this shortcoming, Roe (1989:129) developed a model for the design of selection procedures, which will now be discussed in detail.

3.3.1 The design cycle concept of Roe

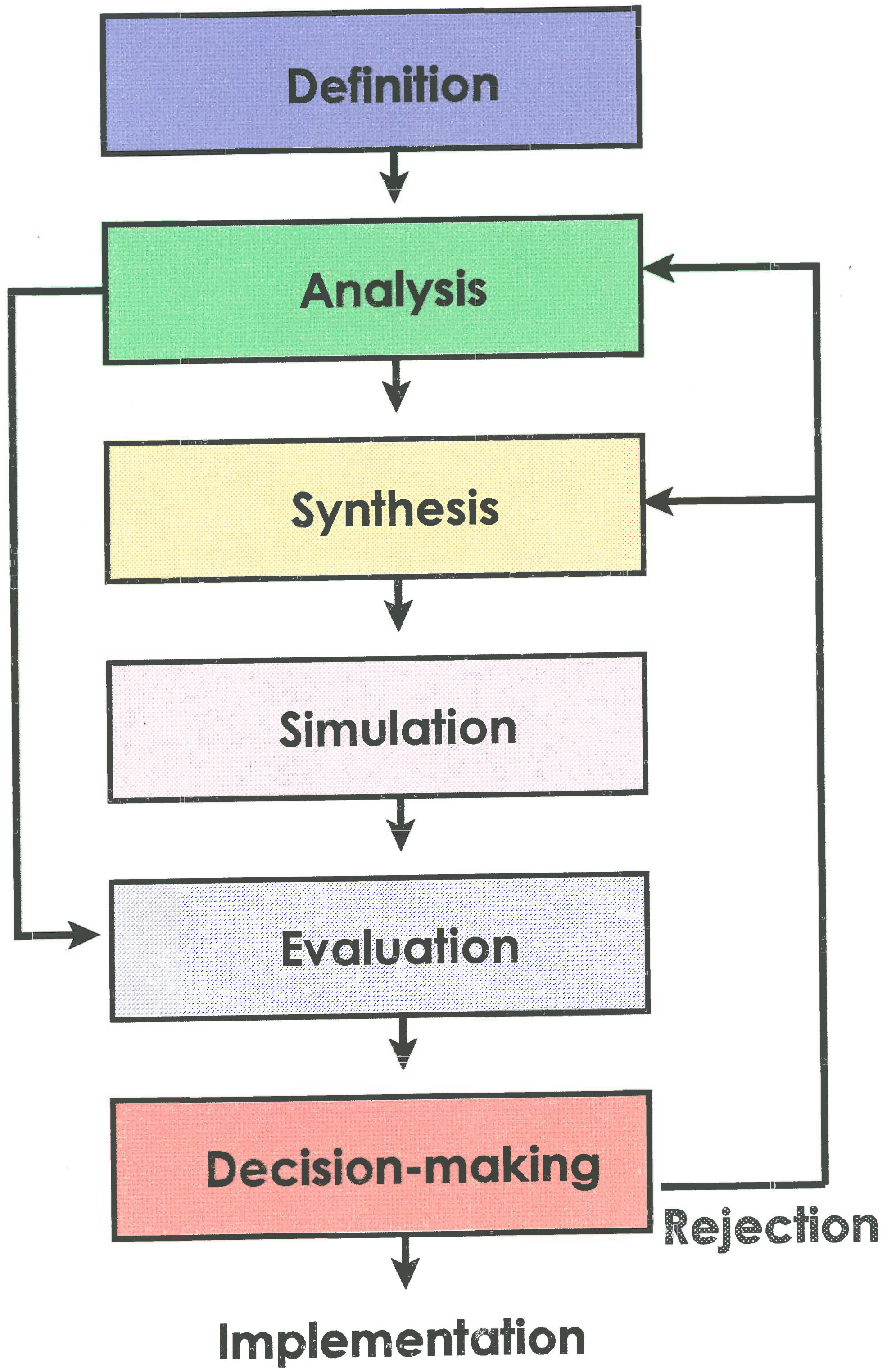
Roe takes the concept of the design cycle from the engineering sciences. It specifies a number of steps that have to be taken within the framework of an iterative process. Although that concept could be used in the design of any product, Roe specifically uses it for designing an integral selection procedure (Roe, 1989:129).

The steps of the design cycle are presented in Figure 3.1 and will be discussed below (Roe, 1989:129-130).

3.3.1.1 Definition

According to Roe, the first step in the design cycle is to define the purpose of the selection procedure and to determine the functions it should fulfil in its particular context. In general, selection procedures are used to collect information, make predictions of performance, evaluate performance and facilitate decisions. These

Figure 3.1: The design cycle model as adapted from Roe (1989:129)



functions can be said to be context-dependent.

3.3.1.2 Analysis

The second step in Roe's model involves the derivation, from the functions to be carried out of requirements that the procedure should meet. These may involve the input data, the prediction, the decision, etc. The constraints influencing the procedure should furthermore be specified, for example, time and resource limitations.

3.3.1.3 Synthesis

The next step is creating a preliminary design or adjusting an existing process in such a way that it fulfils the desired functions, and at the same time, stays within its restrictions. Synthesis is a creative activity, utilising knowledge about people and their behaviour as well as available tools and techniques, to reach an innovative solution for the selection issue at hand. "The result of synthesis is a description of the selection process " (Roe, 1989:130).

Synthesis may include specifications, such as a minimum predictive validity.

3.3.1.4 Simulation

In this step, the operational, predictive and economic properties of the selection procedure are evaluated. Simulation can be carried out on an empirical basis by having experimental runs or by using models such as the Taylor-Russel tables, Curtis and Alf tables or the Cronbach-Gleser formulae for estimating utility.

3.3.1.5 Evaluation

As soon as the properties of the selection procedure are known, their value for the user can be assessed, taking the functions and restrictions into consideration. This step should prove the procedure either satisfactory or unsatisfactory.

3.3.1.6 Decision making

The last step in the designing cycle is to accept the procedure for operational use or to reject it. In the case of rejection, the process may start again at Step three, revising the model in order to improve the previous solution. If errors or insufficiencies in the programme of requirements and limitations show up, the process may return to Step two and the requirements and constraints could then be reformulated.

According to Roe (1989:130), the principle of iteration is typical for the designing process, as a satisfactory solution is usually only found after a number of efforts. The design cycle can be utilised equally well to develop new selection processes and the redesign of existing procedures.

3.3.2 Major functions of the selection procedure

According to Roe (1989:131-137), the following are the most common functions that selection procedures have to fulfil.

3.3.2.1 The information-gathering function

To be able to select an employee for a position in an organisation, specific information regarding both the organisation and the applicant is needed, such as the position and minimum requirements or abilities and traits of the applicant. In the case of developing a selection procedure for admission to higher education, information involving the relevant academic course as well as information on the applicant will have to be accumulated.

The most important choices to be made concerning the information-gathering process are which content area is to be covered and in which format the information will be presented. The outcome of these choices will determine the instruments and techniques to be employed in order to gather the relevant information

3.3.2.2 The prediction function

Within the selection process expectations about future behaviour of applicants are derived from past or present characteristics. Both Roe (1989:133) and Greuter (1989:186) refer to two main predictive principles, namely, the sign approach and the sample approach. The two approaches differ in their underlying epistemological basis and lead to different types of diagnostic processes and the utilisation of different instruments. (Roe, 1989:133.)

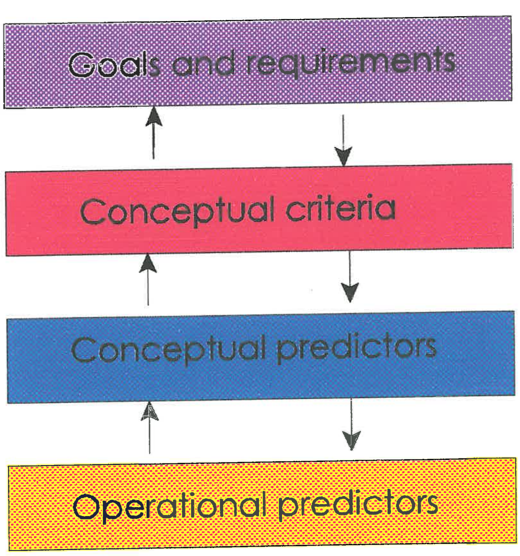
3.3.2.2.1 The sign approach

According to Roe (1989:133), the sign approach is based on the *deductive-nomological* principle. This means that for a given set of people, "when...a certain law states that a relationship exists between a characteristic A and a certain type of behaviour E, one can deduct from this law the prognostic proposition that a person who possesses A will show behaviour E." (Roe, 1989:133.)

Figure 3.2 shows the diagnostic process according to the sign approach, as presented by Roe (1989:134). The process has two phases. The first is an analytical or downward phase, during which diagnostic indicators are identified. First, a description of the organisations goals and requirements are given, whereafter conceptual criteria are specified. These criteria should correspond to relevant dimensions of work behaviour and performance results. Conceptual predictors are chosen on the basis of the differential psychology ("laws" previously referred to). "Finally, the conceptual predictors, which have a well-defined theoretical status, or construct validity, are operationalized by choosing, or developing, certain operational predictors, such as cognitive tests, personality inventories and rating scales." (Roe, 1989:134.)

In the second, or upward, phase of the diagnostic process corresponding to the sign approach, the order is reversed. The first step is to administer operational predictors and the second to make inferences on applicants' traits (conceptual predictors) and work behaviours (conceptual criteria). Finally conclusions are drawn regarding an applicant's predicted success in terms of the organisation's expectations.

Figure 3.2: Diagnostic procedure in the case of the sign approach (Roe, 1989:134)



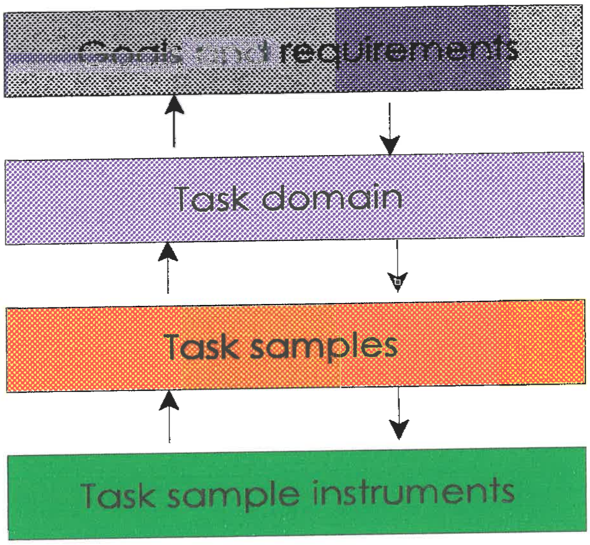
3.3.2.2.2 The sample approach

Greuter (1989:187) states that the sample approach refers to prediction in terms of iconic or analogue models. Roe (1989:133) says that the sample approach -

> "...rests on the principle of generalization: when a person behaves in manner E at a given occasion defined by time and place, it is concluded that he or she will behave identically on other occasions belonging to the same universe."

Figure 3.3 shows a diagnostic process corresponding to the sample approach as presented by Roe (1989:135). The process starts with a description of the organisation's goals and requirements. The second step entails the defining of the domain of tasks corresponding to the relevant position. In the third step a sample is drawn from the domain of tasks set in the second step. This sample is operationalised in a set of activities, such as biographical questionnaires, situational tasks, etc. The results are interpreted in terms of the performance on the task sample, the total task domain and job success.

Figure 3.3: Diagnostic procedure in the case of the sample approach (Roe, 1989:135)



According to Roe (1989:135), the two approaches can be implemented in either a clinical or formalised method and the combination leads to four different forms of prediction, namely:

- *Nomological model*: The model contains a formalised specification of the relationships of one or more predictor variables, operational measures of traits and one or more criterion methods.

- *Domain sampling model:* Content-oriented devises are used to measure past or present performance; scores are generalised in a formal way (for example, statistically) to future performance estimates.

- *Predictor comparison* The scores of applicants on predictor variables are compared in order to find those with the best overall profile. It is assumed that this person's performance will be the best.

- *Criterion analogies*: The work performance of applicants in similar situations are analysed in order to draw analogies. In this way, an idea of future performance is derived from past performance.

3.3.2.3 Decision-making function

The third main function of the selection process, according to Roe (1989:136), is that of decision-making. Although expectations about applicants' future performance play an important role, they do not constitute the end result of the selection procedure. A decision strategy should be generated, which is broadly speaking a way of weighting utilities and probabilities of outcomes and finding the optimal outcome. A formal or informal way of decision-making can be used.

3.3.2.4 Information supplying function

The final main function of selection is the supplying of information about assessment results, expected performance of applicants and decisions taken, to the relevant role players (Roe, 1989:137).

3.3.3 Applying the design cycle to predictive performance models

Greuter (1989:183) says the design cycle as suggested by Roe (1989:129) can be applied to each of the basic functions of the selection process as discussed above. His application of the design cycle to predictive performance models will be discussed in detail here, as it has significance for this study.

According to Greuter (1989:183), prediction models can be seen as special cases of performance models.

> "In a performance model the performance criteria under study are explicitly related to a set of exogenous variables. Exogenous variables can be conceived as determinants of the performance behaviour under study. They serve to predict the future outcomes of performance criteria or they may facilitate the understanding of performance behaviours." (Greuter, 1989:183.)

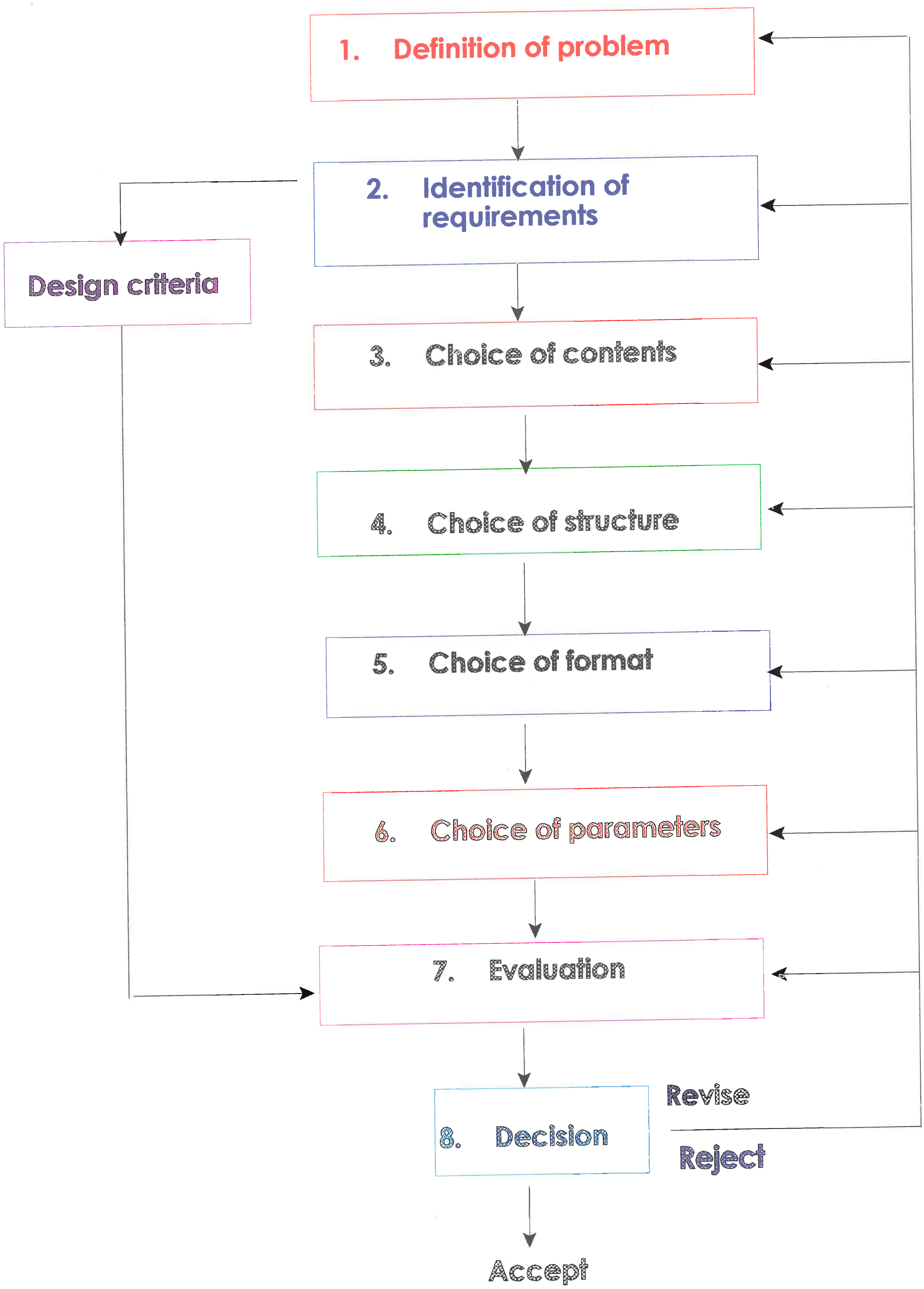
Greuter (1989:184) continues by saying that, in selection, a performance model should contain variables that -

- are relevant to the problem at hand
- can be assessed at the moment of application
- are stable enough to allow predictions over a longer time period
- are streamlinable in order to get an acceptable utility-cost ratio

Greuter (1989:185) defines a prediction model, in the context of personnel selection, as "a model that transforms information about applicants' past or present behaviour into forecast of their future behaviour." Prediction models can vary from statistical regression formulae to job simulations.

Figure 3.4 shows Greuter's application of Roe's design cycle (Greuter, 1989:184), to indicate which steps are to be taken to develop predictive performance models. Each of the steps will be described below.

Figure 3.4: Steps in the development of predictive performance models in accordance to the design cycle model. (Greuter, 1989:184)



3.3.3.1 Steps in developing prediction performance models for selection

3.3.3.1.1 Defining the problem

Performance models aim to predict the success of applicants from the specific characteristics and/or behaviour styles as assessed at the time of their selection. The question asked can thus be restated by formulating predictive performance models that can transform information on past or present applicant characteristics into predictions about future performance.

3.3.3.1.2 Specifying model requirements

Predictive validity constitutes the main model requirement, as predictive performance models should allow for accurate predictions of work performance. Models should furthermore be both practical and economical.

3.3.3.1.3 Specifying model content

Relevant criterion variables should first be specified. Thereafter predictor variables must be identified so as to be able to predict expected performance levels on these specified criterion variables.

3.3.3.1.4 Structure

After the relevant model elements have chosen, the relation between these elements should be specified. The focus should be on relations within and between both sets of variables (i.e. criterion and predictor variables).

3.3.3.1.5 Form

An algebraic function is usually chosen to depict the relationship between criterion and predictor variables in a performance model, for example $Y = f(X)$.

3.3.3.1.6 Estimating parameters

Model parameters are estimated in order further to specify the relationship between criterion and predictor variables. Parameters contain information about the strength of the relationship concerned. They relate to the specification of the model's structure according to which the direction of a relationship is depicted.

3.3.3.1.7 Evaluation of the model

The performance model is evaluated in relation to the model requirements specified at the beginning of the process. Evaluation is primarily directed at determining predictive validity, and the model is assessed in terms of practical and economical considerations.

Model revisions may be necessary as a result of the evaluation process, in which case all of the steps discussed might need to be repeated, until the final results are satisfactory.

3.3.4 Considerations in the design of selection procedure

Roe and Greuter (1991:217) identify four types of considerations that may play a role in the design of a selection process:

⇨ Effectiveness considerations: The appropriateness of the predictive information and the correctness of the decisions taken are taken into account. These considerations influence the selection of, for example, specific psychometric tests or the inclusion of certain questions in the biographical questionnaire.

⇨ Efficiency considerations: The relationship between the overall costs of the selection procedure and the benefits derived from its utilisation is assessed.

⇨ Ethical considerations: Aspects such as the non-intrusion of privacy, right to appeal and fairness of the process are considered. These considerations may for instance result in the setting of different selection ratios for applicants of different cultural backgrounds.

⇨ Managerial considerations: The organisation of the staff members involved in the selection process is evaluated. These aspects may for example, lead to the automation of parts of the process to cut labour costs.

It seems that the technological model suggested by Roe and applied by Greuter, offers an adequate framework for dealing with all of the above mentioned considerations that should be taken into account in the development of a selection procedure. This model will therefore be applied in the execution of this study.

3.4 Redefining "fairness" in selection

According to Arvey and Sacket (1993:172), psychology, for a very long time, treated fairness as a psychometric characteristic of a selection device. They refer to Hunter and Schmidt as being the first theorists to clarify fairness as a philosophical issue, and not as a technical characteristic of a psychometric test.

This view has gained popularity in the ensuing period and it is now reflected in current professional standards. The Society for Industrial and Organizational Psychology states in its official Principles for the Validation and Use of Personnel Selection Procedures that -

> "...fairness is a social rather than a psychometric concept. Its definition depends on what one considers to be fair. Fairness has no single meaning, and, therefore, no single statistical or psychometric definition. Fairness, or lack of fairness, is not a property of a selection procedure, but rather a joint function of the procedure, the job, the population, and how the scores derived from it are used" (1987:18).

Support for the expansion of the definition of fairness is also voiced by the American Educational Research Association, American Psychological Association and the National Council of Measurement when they state that "unlike selection bias, however, fairness is not a technical psychometric term; it is subject to different definitions in different social and political circumstances" (1985:13).

Arvey and Sacket (1993:173) add to the above by describing the selection system content, context, process, and outcome factors, all as potential contributors to fairness.

3.5 Strategies for selection fairness

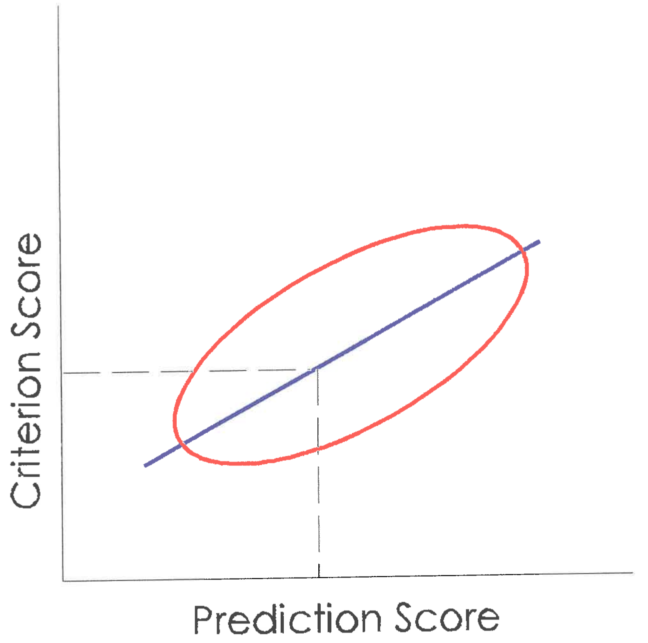
Huysamen (1996:200) states that there are "various strategies or models by means of which selection variables may be combined to reach a decision wether to reject an applicant to an institution of higher education." The most prominent of these will be discussed in order to motivate the decision on the strategy chosen for this study.

3.5.1 The regression models

Huysamen (1996:200) states that regression models choose predictor variables in terms of their ability to predict the criterion of future performance. Taylor and Radford (1986:83) elaborate by saying that the point of departure of the model is the assumption that fairness is achieved in that the applicants with the highest predicted criterion scores are selected, on condition that separate regression equations are used in a case of differential validity between groups.

Cascio (1991:180) states that the comparison of regression slopes and intercepts for various subgroups is widely advocated as a means of investigating selection fairness. He explains that in utilising such a model, predictions are made by drawing a line from any predictor score on the horizontal axis up to the regression line and another line from that intersection across to the vertical axis where the estimated criterion score is indicated (Figure 3.5). Huysamen (1996:200) adds that, if the predictor variables are independent of each other, this equation weighs them proportionally to their respective correlations with the criterion. By implication, poor performance on one criterion can be compensated for by better performance on others. According to Cascio (1991:180,) this model can be used fairly as is if "(O)ver all persons (a) the average error in prediction is zero, and (b) the variance of the errors of prediction is a minimum."

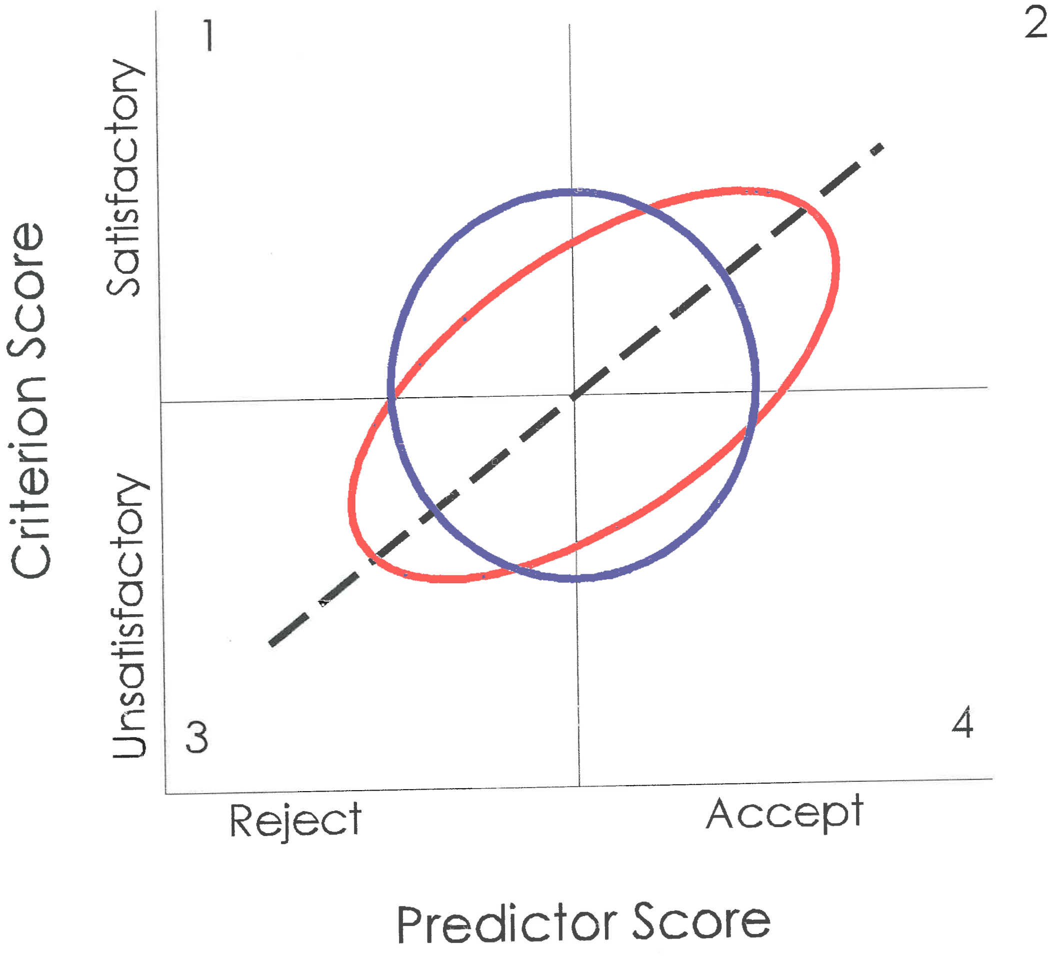
Figure 3.5: Prediction from a single group regression line (Cascio, 1991:180)



Although Taylor and Radford (1986:83) state that no explicit provision is made by the regression model for affirmative action, Huysamen (1996:200) feels that there is nothing that prevents the use of different combinations of predictor variables for different demographic groups. "For each demographic group one may simply use the prediction equation that maximises the prediction of academic performance at the institution involved" (Huysamen, 1996:201). Taylor and Radford (1986:84) suggest that quotas could be incorporated in this model if separate regression lines indicated the need for differential treatment of groups. According to Huysamen (1996:201), a test shows predictive bias if it consistently overpredicts or underpredicts the criterion performance of a specific group. Statistically, this predictive bias is revealed in the regression lines (criterion on predictor) being different for the various groups.

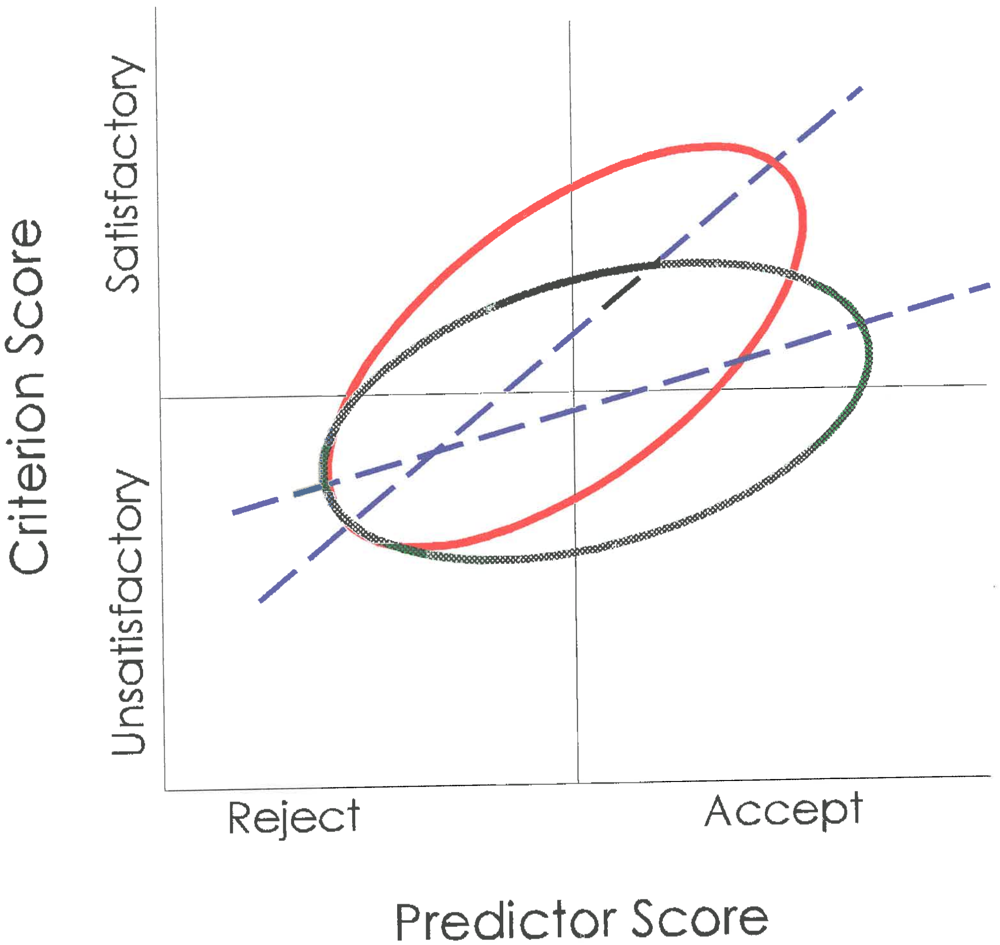
In cases where the distribution forms an **ellipse** (Figure 3.6) a valid prediction is indicated, that is, testees with a low predictor score perform low on the criterion score and those with a high predictor score perform high on the criterion score. Cases falling in quadrants 1 and 3 represent correct decisions, cases in quadrant 4 are erroneous acceptances and those in quadrant 2 are erroneous rejections.

Figure 3.6: A graphic presentation of the distribution of scores in cases of positive validity and zero validity (Schaap, 1995:25)



If the distribution forms a circle (Figure 3.6), there is no predictive validity, as the cases are distributed evenly over all four quadrants and the number of cases correctly predicted does not differ from the number of cases incorrectly predicted.

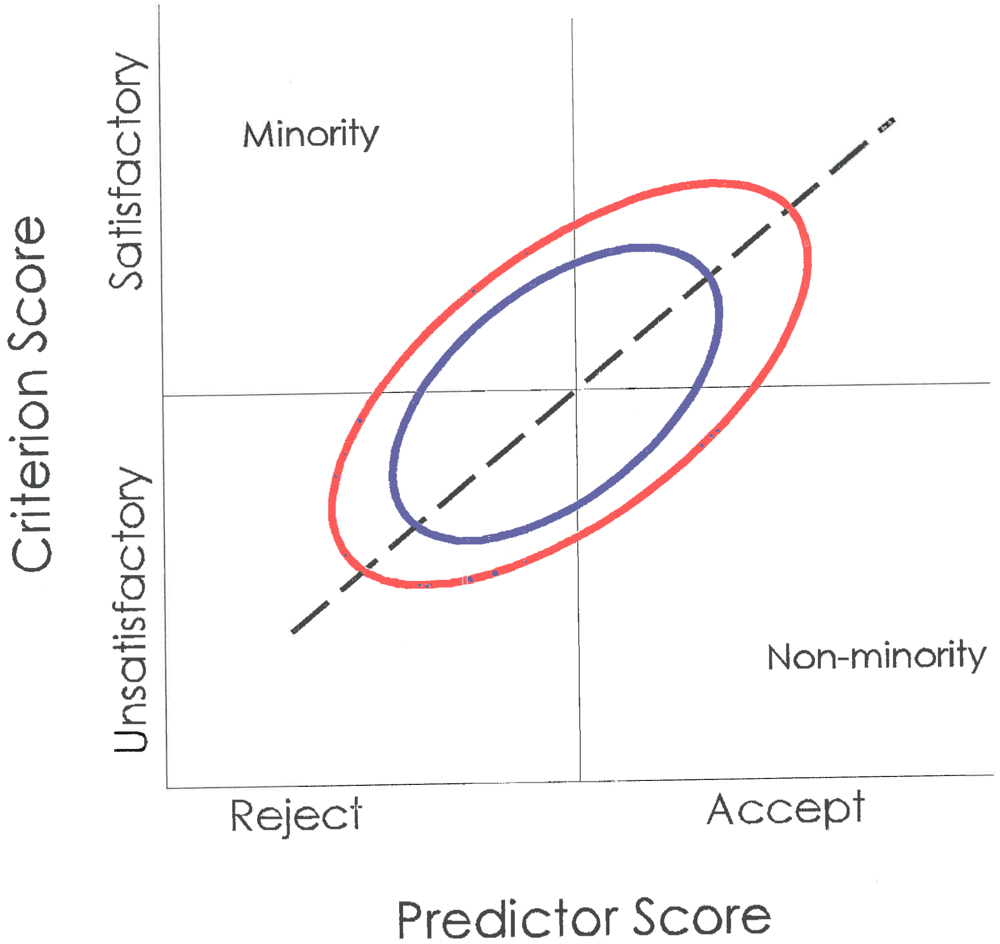
Figure 3.7: Different regression lines and coefficients (Schaap, 1995:25)



If a test shows different regression lines and coefficients for different groups, it has different meanings for different groups, which is indicative of serious test bias (Schaap, 1995:24). This distribution can be seen in Figure 3.7.

Cascio (1991:180) feels that when two subgroups are compared there are three basic possible situations. Figure 3.8 shows the first such possibility - although there are two separate ellipses, one for each subgroup, a single regression line may be cast for both groups. The position of the regression line will be invariant for the two groups, whether calculated together or not, thus there is no need for separate predictions.

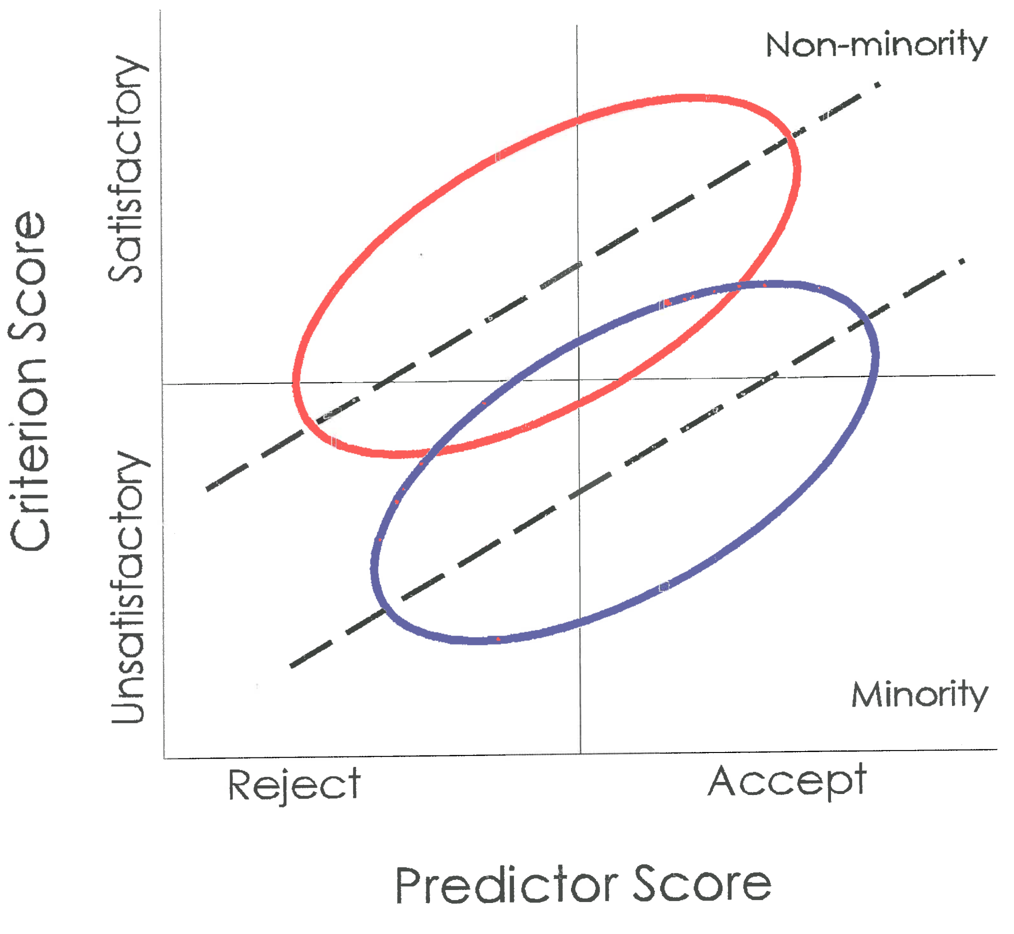
Figure 3.8: Valid predictor with adverse impact (Cascio, 1991:173)



In Figure 3.9, the way in which the position of the regression line is calculated will make a difference. If a single regression line is cast for both groups, criterion scores for the non-minority group will consistently be underpredicted and those of the minority consistently overpredicted. If a single regression line is used, the non-minority group will, in this situation, be affected negatively. Even though the slopes of the two regression lines are parallel, the intercepts are different and the same predictor score has a different predictive meaning in the two groups (Cascio, 1991:181).

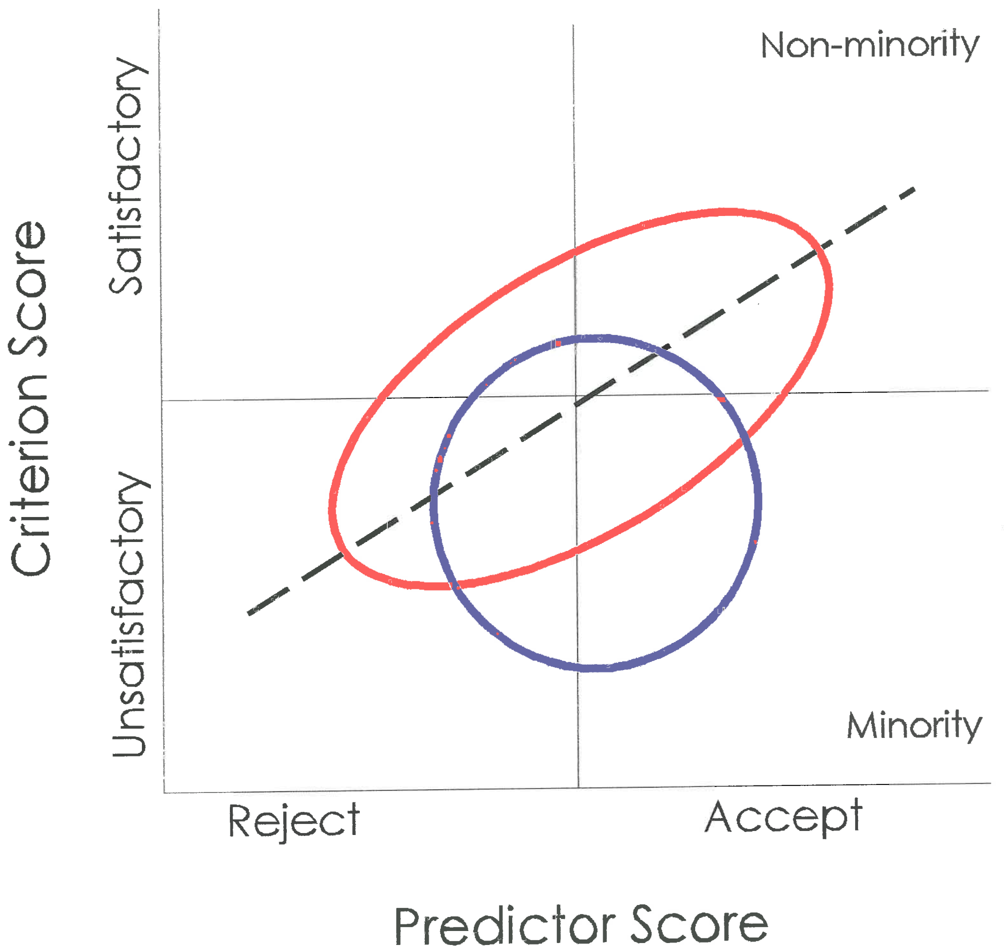
The third possible situation is depicted in Figure 3.10. In this situation the intercepts cross because the slopes of the regression lines are not parallel. The predictor is clearly inappropriate for the minority group in this instance. When the regression lines are not parallel, the predictive validities differ and the intercept differences are meaningless (Cascio, 1991:181).

Figure 3.9: Equal validity, unequal criterion means (Cascio, 1991:174)



Cascio (1991:181) concludes that, despite these potential difficulties, the consensus seems to be that regression models are the appropriate tests for selection bias. The accumulated evidence according to him is clear: "Lower tests scores among minorities are accompanied by a lower job performance, exactly as in the case of the majority" (Cascio, 1991:181).

Figure 3.10: Unequal criterion means and validity only for the non-minority group (Cascio, 1991:175)



3.5.2 Equal risk model

According to the model of Einhorn and Bass (1971:262), a selection model is considered fair if those selected have the same risk of failure, irrespective of their group membership. The distribution of criterion scores about the regression line is examined and separate cut-off scores (minimum predictor scores) should be set if groups have different probabilities of success. The developers of this model emphasise that it is the difference in the standard errors of estimate between subgroups that determines selection fairness. Cascio (1991:182) comments that this method has the same practical implications than the regression model - in groups with equal criterion means, lower predictive validity in one group may decrease the chances of selection for members in that group.

Referring to the ellipse form in Figure 3.6, in this model the probability of success for those selected is the same for different groups when the ratio of the number of cases in quadrant 1 to the number of cases in quadrants 1 and 4 combined is the same for both groups (Cascio, 1991:184).

3.5.3 Constant ratio model

According to Taylor and Radford (1986:84,) the constant ratio model, as proposed by Thorndike, aims at achieving equality of opportunity for different groups. Thorndike (1971: 63-70) proposes that in a fair selection procedure, the cut-off scores are set in such a way that the measure selects the same proportion of minority applicants that would be selected on the basis of the criterion itself or on a perfectly valid selection measure. That is, if 60% of group A and 50% of group B was successful in terms of the criterion, then fair selection would imply a proportional selection of group A relative to group B, at a ratio of 60:50. Cascio (1991:183) comments that Thorndike's model suffers from several disadvantages:

"It is another form of quota setting, it yields overprediction of minority group criterion scores in most situations, and it leads to a greater incidence of placements of individuals into occupational roles for which they are psychologically unsuited. This overprediction for minorities means that certain non-minority applicants will be rejected in favor of minority applicants with lower probabilities of success on the criterion (reverse discrimination). Hence, the overall level of performance will decrease considerably."

Referring again to the ellipse form in Figure 3.6, in the constant ratio model, a selection procedure would be fair as long as the cases in quadrant 2 equal the number of cases in quadrant 4 (Cascio, 1991:184).

3.5.4 Conditional probability model

A fourth model of selection fairness is proposed by Cole (1973:250) and is also aimed at achieving fairness in the relevant proportions of the applicant groups. All applicants who would probably succeed if selected should be guaranteed an equal (or fair) opportunity to be selected, regardless of group membership. Taylor and Radford (1986:84) comment that both this model and the constant ration model "attempt to establish a parity among groups which in effect sets quotas for these groups."

In the conditional probability model, the emphasis is on the proportion of applicants who are above the criterion cut-off. In terms of Figure 3. 6 (the ellipse form), the focus is placed on the ratio of the number of cases in quadrant 1 to the number of cases in quadrants 1 and 2 combined. If this ratio is equal for all subgroups being compared, the selection procedure is considered fair (Cascio, 1991:184).

3.5.5 Modified criterion/Subjective regression model

In this model, Darlington (1971:73) distinguishes between the utilisation of a test to maximise criterion validity and to give preferential treatment to disadvantaged group members. A value judgement is made about the desirability of special selection of members of some group. If special selection is desired, some difference in criterion scores between the groups is adopted, which will yield equally desirable applicants from each group. This is achieved by utilising a formulae by which criteria score *Y* for a minority group is desirable as a score of *Y+k* for non-minorities. Using a variable *C* (with a value of zero for minorities and one for non-minorities), the new criterion will be *Y-kC*.

Cascio (1991:182) calls this method an "esoteric way of setting quotas". He qualifies his remark by explaining that adding or subtracting a constant to the criterion is mathematically equal to adding or subtracting a constant from the predictor score, which in turn is equal to using different cut-off scores for different groups.

In the subjective regression model a selection measure is considered culturally optimal if the proportion of scores fore each group in quadrants 2 and 4 (Figure 3.6) is equal (Cascio, 1991:184).

According to Cascio (1991:185), these five models illustrate that there is more than one reasonable definition for fairness and each definition has some practical and ethical implications which may conflict in some cases. Ultimately, the solution will not be in statistical calculations alone, but also in some values that will need to be weighed. He continues by stating that broader considerations need to receive attention because the validity coefficient by itself falls short of an adequate description of the relative usefulness of the selection decision procedure.

3.6 Decision-making for selection

Selection decisions are concerned with the assignment of individuals to treatments of courses of action which are important to the individual and/or the organisation involved. As the outcome such a decision is not known beforehand to the decision-maker, it has to be predicted on the grounds of available information. This prediction procedure consists of two steps, namely:

- *Measurement*: the collecting of data relevant to the task performance by utilising tests and other assessment tools.

- *Prediction*: the combination of relevant data to ensure a minimisation of predictive error in forecasting task performance (Cascio, 1991:279).

Traditionally, according to Cascio (1991:279), selection programmes had measurement accuracy and predictive efficiency as final goals. This approach has been severely criticised, as it ignores certain external parameters of the situation which have a large impact on the overall usefulness of a selection procedure (Cascio, 1991:292). Two such parameters are the following:

- *Selection ratio*: the ratio of the number of vacancies to the total number of applicants.

- *Base rate*: the proportion of persons judged as successful by using current selection procedures.

Cascio (1991:292) feels that over and above ignoring the aforementioned parameters, the classical validity approach makes unwarranted utility assumptions and also fails to consider the systematic nature of the selection process.

According to contemporary views, measurement accuracy and predictive efficiency are conditions that merely set the stage for the decision problem. Decision theory recognises that the outcomes of predictors are of primary importance and the measurement and prediction are technical components of a system designed to make decisions about the assignment of individuals to tasks and treatments (Cascio, 1991:279). Furthermore, the decision theory forces the decision-maker to consider the usefulness of alternative selection strategies (Cascio, 1991:306).

Within this framework, Cascio (1991:298-303) suggests that the Taylor-Russel, Naylor-Shine and Brogden-Cronbach-Gleser utility models can provide useful planning information to help selectors make better informed decisions.

3.6.1 The Taylor-Russel utility model

Cascio (1991:298) states that if a new predictor is added to a selection programme, the validity referred to by the Taylor-Russel model is based on present employees who have been screened by any other method than the proposed selection procedure. This approach also assumes fixed treatment selection, where individuals are chosen for specific treatments or course of action, rejected individuals are ignored and accepted individuals are divided into successful and unsuccessful categories.

The major shortcoming of this approach, according to Cascio (1991:298), is that the goodness of a predictor is reflected only in terms of the success ratio. Dichotomous classification (successful/unsuccessful) is used to describe criterion performance, and when validity is fixed, success ratio increases as selection ratio decreases. Furthermore, under these circumstances, the success ratio gives an indication that more people are successful, but not how much more successful.

3.6.2 The Naylor-Shine model

This approach assumes a linear relationship between validity and utility which holds at all selection ratios. "...given any arbitrarily defined cutoff on a selection measure, the higher the validity, the greater the increase in average criterion score for the selected group over that observed for the total group" (Cascio, 1991:299).

Therefore the Naylor-Shine model of utility is defined by the increase in average criterion score to be expected from the use of a selection measure with a given validity and selection ratio. The assumption is that the new predictor will be added to the current selection battery and that under these circumstances the validity should be based on the concurrent validity model (Cascio, 1991:299).

This model seems, according to Cascio (1991:299), more applicable than the Taylor-Russel index, because, in many cases, if the selection procedure is valid, an increase in average criterion performance would be expected as the organisation becomes more selective in deciding whom to accept.

3.6.3 The Brogden-Cronbach-Gleser model

According to this model, the effects of the validity coefficient, selection ratio, cost of selection and variability in criterion scores on the utility of fixed treatment selection can be determined. Cascio (1991:300) states that the only assumption necessary to use this model is that the relationship between test scores and task performance is linear, that is, the higher the test score, the higher the task performance and vice versa.

The major strength of this model is that the monetary payoff of selection procedures can be calculated.

3.7 Conclusion

As this study deals with the development of a selection procedure, the application of Roe's design cycle to performance predictor models, as done by Greuter (1989:183), is of specific relevance. Furthermore, the ensuring of fairness in the suggested selection model is equally important and therefore the most effective combination of selection variables should be found in order to make the most accurate decisions possible. In order to measure the said effectiveness, it is important to set valid and reliable criteria by which predicted performance can be evaluated.

The criteria set for the evaluation of the predictive validity of the selection variables will be discussed in Chapter 5.

Chapter 4
Psychometric testing

4.1 Introduction

Traditionally, scholastic achievements were used as selection criteria to gain access to South African higher education institutions. Kotze, Van der Merwe and Nel (1996: 39) state that

> "...mainly because of unequal educational opportunities, (this) placed many students at a severe disadvantage. Although research in South Africa and elsewhere has clearly indicated that matriculation results remain the best predictor of success at tertiary level, it was also found that matriculation results of matriculants from the previous Department of Education and Training and equivalent school systems, particularly at the lower ranges, are an inaccurate reflection of students' academic ability, or potential for success at tertiary level."

Van Aswegen (1997:14) supports the above authors by saying that the implications of the inequalities in the education system are clearly demonstrated by the patterns of access to higher education. He continues by referring to the current crises faced by universities and technikons namely of finding appropriate criteria for undergraduate admission of disadvantaged black students. This is because the difference in the allocation of resources between black and white schools influence the credibility of grades obtained as an accurate reflection of students' academic potential.

From the above, together with the aspects touched on by the White Paper on Higher Education, as discussed in Chapter 2, the need for a fair, reliable and valid instrument to assess the ability of prospective students, in order to see whether they have the potential to be successful in their studies, has become clear.

In order to conceptualise the importance of a fair instrument, it is necessary to understand the current socio-political context of psychometric testing in South Africa. This chapter will first provide the aim of psychometric testing; secondly the current perceptions regarding psychometric testing in South Africa; and will then go on to discuss the nature of psychometric testing.

4.2 The aim of psychometric testing

According to Gregory (1996:34), psychometric testing can be compared to procedures in the physical sciences whereby numbers represent abstract dimensions such as weight or height. Similarly, every psychometric test provides evidence that a person belongs to a specific category and not another. Gregory (1996:35) continues:

> "The implicit assumption of the psychometric viewpoint is that tests measure individual differences in traits and characteristics that exist in some vague sense of the word. In most cases, all people are assumed to possess the trait or characteristic being measured, albeit in different amounts. The purpose of the testing is to estimate the amount of the trait or quality possessed by an individual."

We live in a specialised society and specialised roles should be filled. Cronbach (1990:9) feels that individuals should follow different paths in training in order to fulfil these roles. Psychometric testing, in measuring relevant characteristics, plays an important part in the differentiation of people into specific roles.

4.3 Psychometrics at a crossroad

According to Van Aswegen (1997:6), it is generally agreed that psychometric assessment in South Africa is at a historical crossroad as it tries to free itself from European theories. Nzimande (1995:1) comments that historically disadvantaged groups perceive psychometrics to symbolise the overt and covert racism of South African psychology, as they perceive the development of psychology in this country as closely related to the history of psychometric testing.

The use of traditional psychometric tests for the assessment and prediction of performance of members of disadvantaged groups has been criticised in the past. Van Aswegen (1997:1) states that, among other things, the criticism focuses on the frequent assumption that all people had equal opportunities to comprehend the knowledge and skills the tests require. Hessels (1995:35) makes it clear that this assumption cannot be maintained especially in the case of historically disadvantaged groups with a different linguistic and cultural background. He furthermore feels that these groups are confronted with tests of which the contents are attuned to the dominantly white culture.

Although Taylor (1987:2) acknowledges that some First World countries, for example, the USA, are becoming more and more culturally diverse and are therefore experiencing the same testing problems, he continues by saying that the problems in South Africa are more intense and that psychologists need to devote their attention to addressing those needs by developing appropriate models and instruments. According to Taylor (1987:3), social and political developments could, in future, drive psychological testing from the scene if it does not adjust its role.

At present it is generally known that this prophecy made by Taylor in 1987 has partly been fulfilled and that psychometric testing, particularly for selection and promotion purposes, has been experiencing an enormous amount of pressure, especially from political circles. Schaap, (1997:36) states that criticism is often levelled at psychological instruments, and the discriminating effect of instruments on decisions regarding certain groups and individuals is often emphasised. In the South African, context a negative accent is often placed on the cultural differences as far as they affect validity and the results of measurements and interpretation.

The search for alternatives for psychological tests over the past years has caused a debate of its own. Murphy and Davidshoffer (1994:15) state that: "A critical issue in the debate over psychological tests is wether or not there are alternatives that would be better, fairer, or more economical." The authors continue by referring to the example of the debate on the use of tests in academic admissions in the United States of America. Some critics feel that grade point average (GPA), together with interviews and letters of recommendation, would lead to better decisions than would psychological tests, such as the Scholastic Aptitude Test. Murphy and Davidshoffer (1994:16) comment that the

> "GPA clearly reflects an assessment device that meets our definition of a 'test'. That is, it is based on somewhat systematic observations (typically, a large number of observations from several different courses) under reasonably standard conditions, employing systematic (if not always objective) scoring rules... Both interviews and letters of recommendation fall within the category of behavioral observations...These methods are not based on systematic samples of behavior, they lack standardization and are characterized by highly subjective scoring rules."

In the light of the above it is clear that these methods are less likely to be valid predictors of future performance than well-designed psychological tests. The authors conclude: "Although psychological tests are far from perfect, they represent the best, fairest and most accurate technology available for making many important decisions about individuals."(Murphy & Davidshoffer, 1994:2.)

From the aforegoing it is clear that the discrepancy between the understanding of psychometrics lies in the terms ***differentiation*** and ***discrimination***. Groups in favour of psychometric testing argue that psychometrics categorise people by measuring certain traits and by summing up performance in numbers or classifications (Gregory, 1996:35; Murphy & Davidshoffer, 1994: 2; Cronbach, 1990:9). Groups opposing the use of psychometrics feel that tests assume that all people had the same opportunities to comprehend the knowledge and skills these tests require (Van Aswegen, 1997:1). They argue that psychometric tests put testees from a disadvantaged background in a situation where they are confronted with contents attuned to a white middle-class culture (Hessels, 1995:35).

In an attempt to find a solution to the difficulties facing psychometric testing in South Africa, the idea of assessing learning potential has gained increasing popularity among politicised academics, unionists and personnel practitioners alike. According to Van Aswegen (1997:7), the popularity the concept of learning potential in South Africa can be ascribed to its promise of providing a means of fair assessment, despite unequal educational opportunities. The National Union of Metal Workers (NUMSA) emphasises in its policy document (1992) that tests for learning potential ought to be used. Nzimande (1995:6), highly critical of psychometric testing in general, feels that the assessment and identification of potential might be useful in redressing historical inequalities.

Shirley (1994:1) states:

> "...in order to realise our ambitions of redressing the wrongs of the past without compromise in quality performance, we will have to come to terms with the following, in the cognitive domain:

- ➩ Measures must concentrate on the identification of potential rather than on measures which reflect crystallized competencies or skills if they are to be fair to groups originating from impoverished educational and developmental contexts;

- ➩ The implication is that we should look for modifiability in domain specific skills - assessment of potential which will also be good predictors of performance in actual work settings..."

In the debate on equity and efficiency, Anastasi and Urbina (1997:545) state the other side of the coin when they say that insofar as culture influences behaviour, this effect will and should be detected by psychometric tests. They continue:

> "If we rule out all cultural differentials from a test, we may thereby lower its validity as a measure of the behavior domain it was designed to assess. In that case, the test would fail to provide the kind of information needed to correct the very conditions that impaired performance."

To accommodate the demands placed upon psychometric testing, many test developers have worked towards changing the face of psychometric testing.

To Erasmus (1997:52), this *changed face* implies a transition from testing to assessment when he says that to "test" means to subject a person to a process in which he or she could either fail or pass, whereas assessment should be seen as the mere mental finger printing of a person's potential, with such an assessment only applicable at the point of assessment. He sees potential as a phenomenon which could be manipulated or altered over time through interventions (that is, turning potential into performance), such as training and exposure to a stimulating environment. This presupposes a dynamic approach which recognises the need for assessment and reassessment in order to establish an index of growth.

Schoombee (1992:20) suggests that the development of psychological tests in different parts of the world tends to be more and more scientific and research-orientated. Anastasi and Urbina (1997:27) add that psychometric testing is in a state of rapid change.

As new and revised tests are being published in abundance, it would be worth our while to have a theoretical look at the nature of a scientifically accountable psychometric test.

4.4 The nature of psychometric tests

4.4.1 Definition

A study of the literature reveals many definitions of psychometric or psychological tests.

Sax (1980:13) gives the following definition: "A test may be defined as a task or a series of tasks used to obtain systematic observations presumed to be representative of educational or psychological traits or attributes."

Walsh and Betz (1985:20) say "...a psychological test is a highly refined and systematized version of the ordinary process of observation of ourselves and the people around us."

Cronbach (1990:32) defines a psychological test as "...a systematic procedure for observing a person's behavior and describing it with the aid of a numerical scale or a category system."

Murphy and Davidshofer (1994:3) state that:

"A psychological test is a measurement instrument that has three defining characteristics:

1. A psychological test is a sample of behavior.
2. The sample is obtained under standardized conditions.
3. There are established rules for scoring, or for obtaining quantitative (numeric) information from the behavior sample."

Gregory (1996:33) describes a psychometric test as "a standardized procedure for sampling behavior and describing it with categories or scores. In addition, most tests have norms or standards by which the results can be used to predict other, more important behaviors."

Anastasi and Urbina (1997:4) define a psychological test as an objective and standardised measure of a sample of behaviour.

Smit (1991: 25) gives a comprehensive definition encompassing all the above when he says that "**a psychological test can be described as an objective and standardised measurement of a specific sample of human behaviour. A psychological test therefore measures an aspect of human behaviour from which estimations of the testee's natural abilities and personality characteristics can be made for diagnostic, prognostic or predictive purposes.**"

4.4.2 Standardisation

It is clear from the above definition of psychometric testing, by Smit, that the concept of standardisation is important in the understanding of the nature of psychological testing.

To Walsh and Betz (1983:21) standardisation means that constant procedures and materials are used and that the methods of scoring the test are consequent across all testings. Mehrens and Lehmann (1991:280) add to this definition by commenting that the scoring in a standardised test is usually objective.

Gregory (1996:33) says that a "test is considered to be standardized if the procedures for administrating it are uniform from one examiner and setting to another." Anastasi and Urbina(1997:6) elaborate by placing emphasis on the administering of the test to a large representative sample, known as the standardization sample, in order to establish the norms. According to them this is important because, in most instances an individual's test performance is interpreted by comparing it with that of others.

4.4.3 Objectivity

According to the accepted definition (Smit,1991:25), psychological testing is characterised by objectivity.

Anastasi and Urbina(1997:7) consider a test to be objective if the administration, scoring and interpretation of scores are independent of the subjective judgement of a particular examiner.

Smit (1991:26) quotes various authors in order to explain the concept of objectivity. His conclusion can be summarised as meaning that a test is objective if -

- a testee is not conscious of the way in which his answers can influence the interpretation of the test;
- the correctness of an answer can be determined unambiguously; and
- the test results are independent of the subjective judgement of an examiner.

4.4.4 Validity

There seems to be consensus among writers that validity is the most important characteristic of a psychometric test (Brown, 1983:19; AERA, APA, NCME, 1985:9; Anastasi & Urbina, 1997:8). Validity is said to refer to the extent to which a test measures what it is designed or developed to measure (Brown, 1983:19; Walsh & Betz, 1985:56; Anastasi & Urbina, 1997:8; Kline, 1993:15). A number of theorists add that validity involves the extent to which appropriate and meaningful inferences can be made from test scores and other measurements (Sax, 1980:289; Brown, 1983:98; AERA, APA, NCME, 1985:9; Mehrens & Lehmann, 1991:265; Gregory, 1996:107).

Test validation refers to the process of accumulating evidence to support the inferences made from test scores, therefore the inferences regarding specific uses of a test are validated, not the test itself (AERA, APA, NCME, 1985:9; Gregory, 1996:107). Anastasi and Urbina, (1997:113) support this by saying that the validity of a test cannot be reported in general terms, as it has to be established with reference to the particular use for which the test is being considered. Mehrens and Lehmann(1991:266) stress this argument further by saying that it is only the degree to which the evidence supports the inferences that are made from the test scores that can be labelled valid or invalid.

Traditionally, validity is divided into three categories, namely, content-related validity, criterion-related validity and construct-related validity. (AERA, APA, NCME, 1985:9; Anastasi & Urbina, 1997:114; Murphy & Davidshofer, 1994: 107.)

4.4.4.1 Contend-related validity

According to Gregory (1996:108), content validity is "determined by the degree to which the questions, tasks, or items on a test are representative of the universe of behavior the test was designed to sample." The determination of content validity is usually a judgmental process that provides no quantitative index or measure of validity (Brown, 1983:133).

In selection, inferences based on content validity are concerned with whether or not a test or measurement procedure contains a fair sample of the performance domain it is supposed to represent. If a selection procedure is thus content valid, it will give the applicant a fair chance to demonstrate his or her competence to perform the relevant task (Van Aswegen, 1997:19).

4.4.4.2 Construct-related validity

Construct validity can be described as focussing on the test score as a measure of the psychological characteristic of interest (AERA, APA, NCME, 1985:9). In determining construct validity, testable hypotheses are deduced from theory and data is then collected to test these hypotheses. If the hypotheses are substantiated, the conclusion can be drawn that the test measures the construct at interest (Brown, 1983:139). According to Anastasi and Urbina (1997:126), any data enlightening the nature of the construct under consideration represents appropriate evidence for this validation.

4.4.4.3 Criterion-related validity

Smit (1991: 58) states that criterion-related validity is specifically important in tests used for selection purposes. Therefore it is important for the purpose of this study. A detailed discussion will thus be given.

Criterion-related validity refers to the empirical technique of studying the relationship between test scores and an independent, external measure or criterion (Mehrens & Lehmann, 1991:268). Therefore criterion-related validity indicates the effectiveness of a test or measure in predicting an individual's performance in specific activities (Anastasi & Urbina, 1997:118). It is important to remember that the inferences made from the test scores are being validated and not the test scores per se (Brown, 1983:100).

According to Brown (1983:99), the emphasis in criterion-related validity is fundamentally placed on the criterion performance, as we are interested in the test scores because they predict some important external behaviour. He further states that the idea that tests are used as part of decision-making processes is implicit in the concept of criterion-related validity. "Because the validity of a test is judged by its relationship to the criterion, if the criterion measure does not adequately reflect the desired outcome, the decision-making process will be less effective." (Brown, 1983:100.)

Authors distinguish between two types of criterion-related validity, namely, **predictive validity** and **concurrent validity** (Brown, 1883:99; Walsh & Betz, 1985:58; Anastasi & Urbina, 1997:118; Mehrens & Lehmann, 1991:268; Gregory, 1996:111).

Predictive validity refers to the accuracy with which early test data can be used to estimate criterion performance in the future (AERA, APA, NCME, 1985:11).

Concurrent validity is studied when the predictor and criterion scores are obtained simultaneously and the present status on the test and the present status on the criterion are being observed (Walsh & Betz, 1985:58).

4.4.4.3.1 Selecting the criteria

Brown (1983:101) states that if a test was developed to predict performance in a field, a standard measure of performance must be identified - the criterion. He continues by saying that in most cases, having multiple criteria is better than having one criterion only, as the latter causes valuable information to be neglected.

Latham and Wexley (1981:72) as well as Cascio (1991:68), refer to the debate between theorists on the utilisation of composite versus multiple criteria as measure of performance. The basic argument for composite criteria states that the measurement or criterion should provide an overall yardstick of success for each individual respondent. Latham and Wexley (1981:72) discuss three methods by which composite criteria can be established:

- Each criterion may be weighted equally. This practice assumes that each criterion is equally important for defining overall success on the task.

- The criteria can be subjectively weighted by so called experts.

- The criteria can be weighted in terms of their monetary value for the organization.

Advocates of multiple criteria argue that measures of demonstrably different variables should not be combined (Cascio, 1991:68). As this study is about the validity of a selection battery for an academic programme, multiple criteria will primarily be used. The reason for this decision is that the different subjects which students have to pass differ vastly from each other. Furthermore, none of these subjects are more important than the other, as no qualification can be obtained before all of the prescribed subjects have been passed.

Mehrens and Lehmann (1991:269) comment that, when studying criterion-related validity, both the conceptual and operational (measurement) aspects of the criterion should be examined. The criterion is the global concept of successful performance and this conceptual idea is identified by an operational measure (Brown, 1983:101). In this study, the conceptual criterion will be success in engineering studies and the criterion measure will be the academic marks obtained.

In order to be adequate, criterion measures ought to have certain characteristics. The most important feature of a criterion measure is **relevance**. This means that a criterion measure should reflect the important aspects of the conceptual criterion. (Brown, 1983:102; Mehrens & Lehmann, 1991:269).

Secondly, a criterion measure should be **reliable,** for if it varies from time to time, or from situation to situation it cannot consistently relate to other measures (Brown, 1983:102). Mehrens and Lehmann (1991:269) qualify this statement by saying that as the maximum relationship obtained between two variables is equal to the square root of their respective reliabilities, the reliability of the criterion affects the criterion-related validity as much as the reliability of the predictor.

A third characteristic of criterion measurement is that it should be **free from bias**. Bias is commonly caused by criterion contamination that occurs when the criterion score is influenced by the knowledge of the predictor score (Brown, 1983:102; Mehrens & Lehmann, 1991:269), for example, if a teacher knows the IQ of a pupil it might influence his or her rating of that pupil.

4.4.4.3.2 Validity coefficient

A validity coefficient is a correlation between test score and criterion measurement that gives a single numerical index of test validity (Anastasi & Urbina, 1997:141). The higher the correlation the more accurate can the scores on the criterion be predicted from test scores (Brown, 1983:104).

4.4.5 Reliability

Reliability refers to the extent to which scores obtained by the same person on the same test or a different form of the same test, administered on different occasions, remain consistent (Anastasi & Urbina, 1997:84). In these instances, differences among scores are generally known as errors of measurement and it can thus be said that reliability refers to the degree to which a test score is free from errors of measurement (AERA, APA, NCME, 1985:19). These differences cannot be attributed to errors of measurement if maturation or intervention occurred, or if the inconsistency is relevant to the construct being measured (AERA, APA, NCME, 1985:19).

Theorists like Anastasi and Urbina (1997:91-99), Walsh and Betz (1985:48), Mehrens and Lehmann (1991:254), Kline (1993:5) and Gregory (1996:92) distinguish between the following types of reliability:

- **Test-retest reliability** refers to the extent in which the scores obtained on an identical test, administered at different occasions, are similar.

- **Alternate forms reliability** refers to the degree in which alternative forms of the same test yield similar results.

- **Split-half reliability** is obtained by procedures by which two scores are obtained for each testee by dividing the test in two equivalent halves. This provides a measure of consistency regarding the content sampling of the test (internal consistency).

- **Kuder-Richardson reliability and Coefficient Alpha** are methods by which reliability are determined by utilising a single administration of the test. The methods are based on the consistency of responses throughout the test.

- **Scorer reliability** is determined by calculating the relationship between scores given on the same test by different scores. Most tests provide standardised procedures for administration, and variance in the test scores due to the administrator are minimal. With individual tests, utilised for intensive examination, there is however, evidence of variance due to the administrator. It is important to determine this type of reliability when using subjectively scored instruments in research.

Kline (1993:13) states that "high reliability, both test-retest and internal consistency, is essential for a test to be valid."

4.4.6 Test bias

In the literature many different definitions for test bias are found.

Brown (1983:224) defines test bias as follows:

> "**A test can be considered biased if it differentiates between members of various groups on bases other than the characteristic being measured.** That is, a test is biased if its content, procedures, or use result in a systematic advantage (or disadvantage) to members of certain groups over other groups and if the basis of this differentiation is irrelevant to the test purpose."

Taylor and Radford (1986:80) refer to several authors when making the statement that a test is biased if group differences in performance are too large or too small as compared to the group difference on the criterion measure which is used in predicting performance. They state that test bias may reside in test items, the test as a whole, within subjects, in the tester, in the testing context or in the interaction of these factors with other sources. They are of the opinion that test bias can never be eliminated, but that steps can be taken to minimise the effect of test bias, as its potential sources are known.

According to Rust and Golombok, as referred to by Geldenhuys (1996:20), a test is biased if the testing procedure is unfair to members of a specific group. These groups can be defined on grounds of race, gender, language and cultural differences. They continue by making a distinction between intrinsic and extrinsic test bias:

- **Intrinsic test bias** occurs when there is a difference in the mean performance of different groups and this difference can be ascribed to characteristics of the test itself and not to a difference in the trait being measured.
- **Extrinsic test bias** occurs if differences in test performance between groups can be ascribed to a genuine difference in the characteristic being measured.

Among others, Anastasi and Urbina (1997:164) and Brown (1983:224) oppose the view of Rust and Golombok. Brown states that it is important to realise that his definition does not imply that a test is biased merely because members of different groups perform differently. When a group performs differently on a test and the scores reflect a difference in the characteristic or trait being measured, the test is **not** biased. However, Brown (1983:224) concludes that as soon as a mean and/or distribution difference between groups is found, the possibility of test bias should be investigated.

Anastasi and Urbina (1997:164) state that:

> "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 requirement 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."

Brown (1983:227), as well as Walsh and Betz (1985:379), refers to different types of test bias:

- **Content bias** occurs when the content of test items gives a systematic advantage to a specific group of testees, for example when the test contains questions that are more familiar to one group than another. Another form of content bias can be found in item format and presentation, when for instance pictorial material only depicts white males and never females or blacks.

Murphy and Davidshoffer (1994:286) refer to this type of bias as **cultural bias**, where a group of testees had the opportunity to become familiar with the test content and another group not. The example they use is that of test items that are highly academic in nature; the underlying assumption being that the school environment is more foreign to disadvantaged groups. Furthermore verbal items are more likely to be regarded as biased than nonverbal items - the rationale being that verbal items are likely to be presented in standard language (English, in their example), which more closely resembles the spoken language of the middle-class.

- **Internal structure bias** occurs if the internal or factor structure of a test and/or the behaviour of items in relationship to each other differ across cultural groups. This would imply that the test measures different things across groups.

- **Atmosphere bias** refers to the effects of the testing conditions on test takers' performance. For example the type of motivation elicited, factors related to the tester-testee interaction, and factors in the evaluation and scoring of responses.

- **Prediction/Selection bias** is caused when a test has different predictive validity across groups. Selection bias is examined through the comparison of regression equations and regression lines obtained with the different groups. This type of bias was discussed in detail in Chapter 3.

4.5 Conclusion

As there is ample evidence against the utilisation of traditional tests for the assessment and prediction of performance of members of disadvantaged groups, it was necessary to find an instrument which was proved to be valid and reliable in South African circumstances.

For this study, the Potential Index Batteries, developed by Erasmus and Minnaar, were used. Motivation for the selection of this specific instrument will be provided in Chapter 7.

Chapter 5
Criterion development

5.1 Introduction

According to Roe and Geuter (1991:197), the choice and definition of criteria is a traditional theme for psychologists involved in selection. Cascio (1991:50) supports this statement by saying that adequate and accurate measurement is a fundamental problem in personnel psychology. He continues "...although criteria are sometimes used for predictive purposes and sometimes for evaluative purposes, in both cases they represent that which is important or desirable" (Cascio, 1991:50).

5.2 Defining the concept

Cascio (1991:50) defines ***criteria*** as follows: "Criteria are operational statements of goals or desired outcomes." This definition, according to him, is applicable both when the criteria are used for predictive and for evaluative purposes. Criteria should always represent that which is important and desirable.

Pieters (1996: 297) refers to criteria as the specific behaviour, work process and/or outcomes that should be evaluated.

5.3 Criterion problem

According to Leap and Crino (1993:337), the ***criterion problem*** is created by an inadequate performance evaluation system used in a criterion-related study. This is caused by validating the selection device against performance information (criteria) that does not represent true levels of performance. Under these circumstances, the validity coefficient will be inaccurate and undefendable.

5.4 Criterion contamination

Anastasi and Urbina (1997:119) refer to the concept of ***criterion contamination*** as a possible source of error in the validation of selection procedures. According to them, criterion contamination occurs when the scores obtained on predictors themselves influence individuals' criterion status. For example, if a lecturer knows that a specific student has performed poorly on the potential assessment test, such knowledge might influence the grades given to that student. This would obviously raise the correlation between predictor scores and criterion scores in an artificial manner. It is therefore suggested that predictor scores are kept strictly confidential and that no person participating in the assignment of criteria ratings have any knowledge of the testee's assessment scores.

5.5 Dimensionality of criteria

According to Cascio (1991:53), operational measures of criteria vary along various dimensions of which the most important are discussed below.

5.5.1 Temporal dimensionality

Cascio(1991:54) states that the optimum time for criterion measurement varies from situation to situation and that conclusions might be influenced on when the criterion measurement was taken.

As far as this study is concerned the implication will be that academic results taken early in a semester might differ largely from academic results taken later in the semester, due to various factors. Failure to consider the temporal dimension of criteria may thus lead to misinterpretations.

Cascio (1991:55) continues by identifying two special cases of temporal dimensionality, namely:

➢ **Static dimensionality**

The nature of any task performance, if observed at any specific time, is multidimensional. That implies that a number of independent skills are involved in the performance of the task at hand. Leap and Crino (1993:341) support this argument by saying that there is no universal set of performance criteria for all tasks or jobs. If the criteria are not relevant to a specific task, they should not be examined.

In criterion research, the static dimensions are those typically investigated. A "photograph" of performance is taken at a single point in time, usually involving a single criterion, and the assumption is made that the performance of the employee or student has been "captured". Inevitably, this is not always the case.

➢ **Dynamic dimensionality**

As employees (and in the case of this study, students) develop their potential and abilities, the dimensions of performance that seemed important and valid early in their careers may change and seem irrelevant to their performance at a later stage. Therefore it could be said that criteria may be "dynamic" - changing in importance over time.

Dynamic criteria might assume one of three possible forms (Cascio, 1991:56):

a. changes in average group performance over time;
b. changes in validity over time;
c. changes in the rank-ordering of scores on the criterion over time.

Changes in the validity over time, as seen by Cascio (1991:56), are of importance for this study.

Two possible explanations for the changing of validities over time has been given:

i. Changing task model

This model suggests that "while the relative amounts of ability possessed by individuals remain stable over time, criteria for effective performance might change in importance" (Cascio, 1991:56).

As the demands of the labour market change, especially the rapidly changing technological demands, the training of engineering students has to be adapted. This causes the desired performance of these students to change over time and therefore the validity of criteria (and predictors) will change.

ii. Changing subject model

The second model suggests that validities might fluctuate because an individual's level of ability may change over time, even though specific abilities required for effective performance may remain constant.

Referring to the issues of admission to higher education, raised in Chapter 2, the relevance of this model to the study at hand is clear. As admission to higher education has been made more accessible, the profile of the student undergoing training has changed drastically. Due to historical inequality, a large proportion of current students lack certain academic skills that are necessary to achieve success in a higher education environment. This has a definite influence on the validity of predictors and criteria. It was, in fact, this situation which motivated this study.

5.6 Essentials of criteria development

Cascio (1991:57) refers to research done over 40 years ago by Stuit and Wilson, which demonstrated that continuing attention to the development of better performance measures leads to better prediction of performance.

He states that criteria should be developed and examined before predictors are selected to predict the criteria performance: "...if we use predictors with no criteria, we will never know whether or not we are selecting those individuals who are most likely to succeed" (Cascio, 1991:57).

He further identifies the following four basic problems which should be dealt with before human performance can be studied:

5.6.1 Reliability of Performance

This essential issue in personnel research refers, in this context, to the consistency or stability of job or task performance over time. According to Cascio (1991:58), the assumption of performance reliability is implicit in all predictive studies.

Cascio (1991: 58) summarises research done over 30 years of personnel psychology that shows that the reliability of performance varies greatly - individual output was found to be erratic and highly inconsistent. This could lead to misinterpretation in research if the researcher happened to select a period of unusually high or low performance.

Cascio (1991:58) suggests that measures could be employed to compensate for the above inconsistency. One such measure could be to aggregate performance and thereby cancelling out or minimising incidental factors.

The implication for this study would be that final pass marks (a combination of marks obtained throughout the semester) would be used as performance measurements and not, for instance, the results of a single test series.

The main problem experienced when working with a set of final marks was that no mark was available for those students who had decided not to continue with a subject until at least predicate day.

5.6.2 Reliability of job performance observations

This aspect is of cardinal importance in prediction, as all evaluations of performance depend on an observation of some kind, and different methods of observation might lead to different conclusions (Cascio, 1991:59).

Since academic results will be used as criteria in this study, this issue might be addressed by the fact that all students are examined at the same time, by the same paper, scored by the same lecturer. The observations (marks) should therefore be consistent and comparable.

5.6.3 Dimensionality of job performance

Cascio (1991:60) states that research reveals that a variety of predictors are generally used in personnel decisions, but that the majority of studies use a single or global measure of the job or task performance. He asks whether it is meaningful or realistic to reduce performance measurement to a single measure.

This problem is less relevant for this study, as academic achievement as expected from the selected students is not such a complex task that it needs more than academic performance as criteria measurement.

5.6.4 Performance and situational characteristics

According to Cascio (1991:60), this issue deals with the influence of environmental factors and conditions on individual levels of performance. In research it is necessary that attention be given to the possible moderating effects of variables other than those measured by predictors.

5.7 Steps in the development of criterion

Cascio (1991:62) refers to a five-step procedure for criterion development outlined by Guion, namely:

1. Analysis of job or task and/or organizational needs.
2. Development of measures of actual behaviour relative to expected behaviour, as identified in Step 1.
3. Identification of criterion dimensions underlying measures developed in Step 2.
4. Development of reliable measures, each with high construct validity of the identified elements.
5. Determination of the predictive validity of each independent predictor for each one of the criterion measurements.

Smith and Robertson (1986:50-51) identifies three steps in the development of criteria, namely:

1. Analysis of job or task which describes expected behaviour.
2. Analysis of expected outcomes.
3. Analysis of individual's contribution to the achievement of organizational goals.

The last step is not always practical, since it may be difficult to gather this data in time for selection decisions.

5.8 Evaluating criteria

Three concerns should be taken into consideration when criteria are evaluated (Cascio, 1991:63).

5.8.1 Relevance

Any criterion should be logically related to the conceptual criterion (that is, desirable outcomes based on the more general purposes or aims of the organisation). It is therefore important that the conceptual criterion is set clearly.

This point is also emphasised by the American Psychological Association (1985:60): "The rationale for criterion relevance should be made explicit. It should include a description of the job in question and of the judgements used to determine relevance."

Cascio (1991:64) continues by saying that if an important aspect of job or task performance is not assessed, an additional criterion measurement is required, regardless of how many criteria are already used.

5.8.2 Sensitivity or discriminability

A criterion measure, in order to be useful, must be capable of discriminating between effective and ineffective employees (Cascio, 1991:64).

5.8.3 Practicality

Record-keeping and data collections cannot become impractical or interfere with ongoing operations, and therefore personnel researchers should keep criterion measures as relevant and practical as possible (Cascio, 1991:64).

The set of criteria selected for this validation study consists of all academic subjects taken by Civil and Mechanical Engineering Technology students at Technikon Pretoria, for the period 1996 - 1999.

As the general purpose of Technikon Pretoria, as institution for higher education, is to deliver qualified professionals to the labour market, academic success should surely be seen as a relevant criterion. Furthermore, academic performance discriminates and distinguishes clearly between successful and unsuccessful students. Academic records are kept for every student registered at Technikon Pretoria, thus the criteria remain practical as well.

5.9 Conclusion

In order to validate a selection battery for the selection of first-year students, the development of criteria and criterion measures is essential.

In the case of this study, the desired outcome is academic achievement, which will be measured by means of academic performance. In order to have criterion measures that are as valid and reliable as possible, academic performance will be aggregated to compensate for incidental factors. Care will also be taken that, as far as possible, all results used were obtained on the same tests, scored by the same lecturer.

It could be argued that academic performance is not the only relevant measurement when it comes to the selection of students. Issues such as performance on the actual job could also be regarded as relevant. The counter argument could be that a student who will not be successful in the job market should not pass the course.

As this study was initiated to address the specific problem of admitting students to higher education, academic success was deemed criterion enough. Future studies could focus on the predictive validity of the suggested selection battery used in this study for performance in the labour market.

Criteria measurements used in similar studies will be discussed in Chapter 6.

Chapter 6

Predictors of academic success

6.1 Introduction

Zaaiman (1998:61) states that the prediction of future behaviour is an important element of selection. She adds that the behaviour of each individual is influenced by a variety of interacting factors, and that, for selection purposes, one should identify those factors which can be measured in practice and which make significant contributions to the prediction of eventual academic performance.

In this chapter, attention will be given to such possible predictors, as identified and examined by previous research, namely, scholastic performance, learning ability, general cognitive ability, specific aptitudes and biographic factors such as gender and age. Subsequently, because of the focus of this study, possible predictors for success in engineering technician courses, as found in literature, will be discussed.

6.2 Scholastic performance

According to Zaaiman (1998:61), selection for higher education is internationally based mainly on academic criteria, where academic achievement is usually represented in the form of final school marks. This remark is supported by other researchers such as Kotze, Van der Merwe and Nel (1996:39) when they say that matriculation results have internationally been proved to be the best predictor of academic success at higher education level.

The best known ways of quantifying matriculation symbols are a conversion to the so called M-score and the allocation of a Swedish Formula score. Table 6.1 illustrates the different approaches to quantification.

Table 6.1: Quantification of Matriculation Symbols

Swedish Formula		Matriculation Symbol	M-scores	
HG	SG		HG	SG
8	6	A	5	4
7	5	B	4	3
6	4	C	3	2
5	3	D	2	1
4	2	E	1	0
3	1	F	0	0

Over the past two decades a number of research findings showing the relationship between scholastic achievement and first-year higher education success have been published.

Stoker, Engelbrecht, Crowther, Du Toit and Herbst (1985:7) reported on twelve South African studies between 1957 and 1977 that " they all found that success at school did to a fairly large degree extend to achievement at university, particularly first year achievement". Bokhost, Foster and Lea (1992:59), however, comment that most of those studies were conducted at white, mainly Afrikaans-speaking universities.

In the study commissioned by the Committee of University Principals and conducted by the HSRC (Stoker et al, 1985), a number of factors were included, namely, age, school aggregate, gender, population group, examining body, home language and university. From this study it emerged that school aggregate was the best single predictor of university performance; specifically in the BA study field for both language and non-language directions.

In the science-related fields of study, research has also indicated that matriculation results have a strong relationship with university performance. In the case of the University of Cape Town, Moran (1987:2) expressed the following: "The evidence is clear and unequivocal. Matriculation point scores were a surprisingly excellent indicator of subsequent academic success for our first-year intake of students."

In a study performed by Bokhorst, Foster and Lea (1992:64) on factors affecting first-year students in Psychology at the University of Cape Town, it was found that the matriculation aggregate had an extremely high predictive effect, appearing as an almost perfect positive correlation. These results should however be interpreted with some reservation, as the number of non-white respondents was very low. Furthermore, significant effects for population group, language and matriculation authority were found, all pointing to a black disadvantage (Bokhorst et al, 1992:64).

At the Rand Afrikaans University, Zietsman (1985:22) found that matric symbols, weighted according to the Swedish Formula, were good predictors of academic success in Chemistry and Botany. Snyman (1987:9), at the same institution examined admission requirements for Law studies. He suggested a minimum of a higher grade D-symbol for first and second language in matric. Jacobs (1987:33) deduces from the above that students who performed well at school level have developed a study attitude and ability to work, which enable such a student to perform well at higher education level.

On the opposite side, many researchers "have expressed grave misgivings about the reliability and validity of school performance as predictors, particularly in the case of black students" (Bokhorst, Forster and Lea, 1992:60). Among these are the Committee of University Principals (1987: 79) who noted the discrepancies within the differentiated educational systems and concluded that school examination results are not always a reliable reflection of a student's academic potential.

Kotze et al (1996:39) states that it was found that matriculation results from the previous Department of Education and Training and equivalent school systems, particular at the lower ranges, are an inaccurate reflection of student's academic ability. Zaaiman (1998:62) supports this by saying that school results are seen as being greatly dependant on educational opportunity and therefore refer more to the opportunity to learn than the ability to learn.

6.3 Learning potential

According to Kotze et al (1996:40) higher education institutions in South Africa have responded in different ways to the issue of alternative selection. One such framework that has been developed is that of the determination of learning potential. According to the aforementioned authors, this paradigm seeks to assess the ability of an individual to respond to actual learning challenges, by the inclusion of a teaching/learning element in the selection procedure.

Zaaiman (1998:64) adds that a dynamic testing framework usually consists of a test-teach-test situation and is aimed at evaluating the student's ability to benefit from instruction. The pretest provides data on the student's current level of functioning on a given task. Instruction is then given in order to teach the student the necessary problem-solving skills. The student's level of functioning is assessed during a post-test session. An improvement in the post-test result is seen as an indication that learning has occurred.

According to Zaaiman (1998:64), dynamic testing instruments are often seen as more fair to use for selection in situations of unequal educational opportunity, as the intervention phase gives the student who has not had the opportunity to develop his or her academic potential a better chance to achieve a fair result.

In studies conducted at the Rand Afrikaans University (RAU) and the Potchefstroom University for Christian Higher Education (PU for CHE), Kotze et al (1996:51) endeavoured to determine the extent to which learning potential could be used to predict academic performance in first-year higher education studies. For the purpose of the investigation, the **Ability, Processing of Information and Learning Battery (APIL)** as developed by Taylor(1991), was used. According to Kotze et al (1996:63), the aim of this test is not to test the student's previous learning experience, but to determine the extent to which the individual has the capacity to acquire new concepts and skills of a cognitive nature, to process the information and apply it.

At the RAU the academic results of 2 336 students were correlated, by means of the Pearson product-moment coefficient, with the APIL scores obtained. Significant correlations were found between the total APIL scores and the academic performance of students in all faculties included in this study.

However, at the PU for CHE, a multiple regression analysis was performed on the APIL data, with first-year academic results as criterion predictor. The APIL was not found to make a meaningful contribution to the explained variance in the academic performance of first-year students (Kotze et al, 1996:79).

This finding is supported by Zaaiman (1998:65) who states that research results have not shown a significantly higher predictive validity for learning potential tests than for previous academic performance. She adds that a dynamic selection procedure is also not practical when a large number of applicants have to be evaluated, as it is resource intensive and time consuming.

6.4 General cognitive ability and specific aptitude

The concept of cognitive or mental ability has been central to the development of psychology over the past century. Despite this, the concept is not well defined and controversy surrounding it is rife (Van Aswegen, 1997:23). According to Zaaiman (1998:62), ability is usually defined "in general terms that refer to a wide range of mental capacities affecting all mental operations." She also states that, in the context of selection, ability tests usually refer to the traditional intelligence type tests.

Thorndike (1986:332) states that ability testing, in its beginnings, focused on providing a measure of general cognitive functioning. Zaaiman (1998:62) adds that ability tests are non-subject specific and gives the following examples of ability tests:

- verbal
- numerical
- pattern recognition
- spatial reasoning

The usefulness of assessing general ability as predictor of performance has widely been debated. In 1986, the Journal of Vocational Behavior (Gottfredson, 1986) devoted an entire issue to the topic of general ability and its role in the prediction of task performance. In this publication a number of distinguished psychologists address the question of whether or not there is an association between mental ability and job performance. The conclusion that is reached in this volume is that general cognitive ability (or "g" as it is also referred to) emerges as the single most useful task performance predictor.

According to Zaaiman (1998:62), the opposite seems to be true within the academic sphere. She states that a consistent research finding is that previous academic achievement has the highest predictive validity, with traditional aptitude tests next and ability measures having the lowest predictive validity. A possible explanation for this could be that ability tests measure concepts that are only hypothetically related to academic achievement and are thus further away from the task to be predicted than either achievement or scholastic aptitude tests. Furthermore, academic performance usually constitutes the first predictor, but due to its high correlation with ability tests, the ability tests are not included in the prediction model. Suppressor variables are then included as second and third predictors in the prediction model. Zaaiman (1998:62) continues that one could expect that a minimum level of prior academic achievement would be necessary for a student to benefit from an academic programme, regardless of inherent ability.

From the early concept of general ability a movement towards more specialised ability tests developed and, in time, the rationale and procedures for combining the results of such tests, using the statistics of multiple regression to create task-tailored batteries were put forth (Thorndike, 1986:332). " Following up on the theoretical emphasis on specialized abilities, tests of such special abilities began to multiply and aptitude test batteries began to replace general ability tests" (Thorndike, 1986:333).

As early as 1937 Bingham defined aptitude as "...a condition or set of characteristics regarded as symptomatic of an individual's ability to acquire with training some (usually specific) knowledge, skill or set of responses such as the ability to speak a language, produce music etc" (1937:16).

Zaaiman (1998:61) says that scholastic aptitude tests are designed to evaluate subject-related skills, comprehension and insight, utilising non-routine problems which require little subject-specific knowledge.

Contradicting research findings as far as scholastic performance is concerned have been reported by researchers looking into the predictive validity of aptitude tests. Skuy, Zolezzi, Mentis, Fridjhorn and Cockcroft (1996:113) found that the **Patterns Relation Test** of Barker, which is a test of inductive reasoning and reasoning by analogy, predicted performance in Accounting for a group of students in the Faculty of Commercial Science at the Witwatersrand University, significantly ($R^2=0.38$; $p<0.05$).

Pienaar (1991:119) concluded that specific subtests of the Academic Aptitude Test (developed and standardised by the Human Sciences Research Council) differentiate between successful and unsuccessful learners at higher education level. He investigated the predictive validity of the Non-Verbal Reasoning, Verbal Reasoning and Reading Comprehension subtests and found that the academically successful students performed, on average, significantly better on all of the mentioned subtests than the academically unsuccessful students.

In a study conducted at the RAU, the Senior Aptitude Test (developed and standardized by the Human Sciences Research Council) was used to determine the profile of a potentially successful student in social work (Van Zyl, Terblanche & Jacobs, 1992:33). In this study, no significant difference could be found between the Senior Aptitude Test performance of successful students and that of unsuccessful students.

Kotze (1994:64) refers to various studies regarding the predictive validity of the SAT in academic environments, which all conclude that the SAT does not differentiate significantly between potentially successful and at risk candidates.

Samkin (1996:118) reports on a study conducted at the University of Durban-Westville where the performance of first-year Accounting students were predicted, using final school marks as well as subtests of the Academic Aptitude Test (Human Sciences Research Council, 1974). In this study it was revealed that the results of the English Reading and Number Comprehension subtests combined explained 14.4% of the variance of the final Accounting result for the total group of students. "For HED students both English Comprehension and Number Comprehension are individual reliable predictors of success in the first year. However, the results of English Comprehension and Number Comprehension cannot be used as reliable predictors of success for DET students" (Samkin, 1996:118).

6.5 Biographical factors

Various references can be found to the influence of biographical factors such as age and gender on the performance of students on higher education level are found in the available literature.

In this regard, Zaaiman (1998:67) states that gender inequality with respect to access to higher education is of international concern. Brusselmans-Dehairs and Henry (1994:353) wrote the introduction to a volume of the International Journal of Educational Research which is dedicated to the gender differences in mathematical abilities. There is general consensus that boys generally perform better than girls in Mathematics and Science at high school level.

As far as student age is concerned, Zaaiman (1998:68) declares that this might be a factor of importance in the South African context, as many pupils take longer than the norm to complete their schooling and are thus older than 18 or 19 years when applying for admission to higher education. In 1993, according to Zaaiman (1998:68), 37% of the total number of black candidates for the matriculation exam was 22 years and older. She refers to a trend in the UK where older applicants usually come from underrepresented groups in terms of ethnicity and social class. "In the UK the typical mature applicant is aged 21-25, with lower A-level qualifications often gained over several attempts" (Zaaiman,1998:68).

As this study is concerned with the development of a **selection battery** for first-year engineering technician students, biographical variables such as age and gender will not be included since it would be unconstitutional to deny applicants access to further studies on the grounds of their age and/or gender.

It can be seen from this overview that a large number of factors could influence academic performance. All of these cannot be tested or practically implemented in the process of selection, and not all will add to the fairness or predictive validity of selection procedures, but one of the aims of selection research is to identify the best possible combinations of predictors for future performance. Such combinations, specifically developed for the prediction of performance by engineering and engineering technology students will subsequently be discussed.

6.6 Predictors of success in engineering and engineering technician courses

In Chapter 2 reference was made to the statement of the Department of Education in its White Paper on Higher Education (1997:8) that there exists a mismatch between the output of higher education and the needs of a modernising economy. A particular shortage in fields of science, engineering, technology and commerce is mentioned.

To higher education circles this remark came as no surprise. As early as in 1983 Smit(1983:1) stated that South Africa experienced general shortages in high-level artisans and acute shortages in engineers and engineering technicians. He added that in the midst of these shortages attrition rates, especially during the first year of training, remained alarmingly high.

In the sixteen years that have followed, research into the predictors of academic success in engineering and engineering technician courses continued. Those important for this study will be highlighted here.

In a study commissioned by Technikon Pretoria, Smit (1983:62) used a sample of 324 engineering technician students in order to determine predictors for academic success in Electrical, Mechanical and Civil Engineering technology courses. He examined the role played by scholastic performance in Mathematics, Science and Afrikaans (at that stage the medium of instruction at Technikon Pretoria). He further included scores on all of the subtests of the Senior Aptitude Test (SAT) as well as biographical information gathered by means of a questionnaire.

He found that performance in Afrikaans and Mathematics at school level played a prominent role in the prediction of success in technician training (keeping in mind that Afrikaans was the medium of instruction at that stage). Even though performance on subtests of the SAT contributed to a higher predictive validity it was not statistically significant. Non of the biographical factors (such as age, marital status, sex and home language) played a meaningful role in the prediction of students' academic success.

Fourie (1988:77) looked at the scholastic performance of first-year engineering students at the Rand Afrikaans University as predictors of their success at higher education level. Performance in Mathematics was found to be an important predictor of success in that particular group. It was expected by the researcher that mother-tongue performance at matric level would play a significant role in the prediction of higher education success. However, the results of the study showed that the performance in the home language in the final school exams did not predict the academic performance of first-year engineering students at the RAU. It further seemed that the Standard 9 final marks were a more accurate predictor of tertiary success than final matriculation results. It has to be mentioned that this sample was totally homogeneous and that the findings can therefore not be generalised.

Rutherford and Watson (1990:354) stated that selection procedures at the University of the Witwatersrand needed redesigning, since the use of final matriculation marks as sole selection criteria for students in the science and technology fields was unfair and ineffective. They wanted to develop a selection procedure that was acceptable to all, and at the same time admitted only students with a reasonable chance of graduating.

To achieve these objectives, it was deemed necessary to supplement matric results with further information about the candidates.

> "It was felt that strictly content tests, while standardizing achievement measures across the entire spectrum of applicants would further disadvantage the disadvantaged and so measures of ability which had been found elsewhere to be relevant to success in the sciences were sought" (Rutherford & Watson, 1990:355).

They identified the following factors as occurring frequently in relevant literature, and measures thereof was included in the testing battery :

- Piagetian development level
- Spatial ability
- Maths and English ability

The Science Aptitude Test (Atlink, 1987), on close scrutiny, revealed many Piagetian-type questions - the principles of the conservation of mass and volume are, for example tested. Proportional reasoning is furthermore necessary to answer many of the questions (Rutherford & Watson, 1990:356).

The authors stated that, although the test was not purely a Piagetian questionnaire and did not allow deductions regarding the validity of Piagetian development levels in the selection for science courses, the fact that it contained Piagetian concepts that had been linked to performance in science gave it construct validity.

The second factor identified by Rutherford and Watson (1990:356) as being relevant to success in studies in the sciences was spatial ability. The Rotate and Flip Test, developed by the NIPR as a test of spatial visualisation, was selected to form part of the selection battery. In this test, the subject is provided with sections of a figure and has to manipulate them mentally by means of rotating and flipping them to form a given whole (Rutherford & Watson, 1990:356).

The third factor included into the selection battery by Rutherford and Watson (1990:356) was that of Mathematical and English ability - generally linked to success in the sciences. The authors added that competence in English was an obvious prerequisite to success in this particular course, as it was taught in English.

The Mathematics and English factors were combined in the selection of the third test of the selection battery, namely, the Embedded Problems Test, also designed by the NIPR (Rutherford & Watson, 1990:356). The researchers criticised their own decision by saying that the Embedded Problems Test was not an ideal test as "the scoring does not allow credit for disembedding ability, English ability and Mathematics ability to be calculated separately, as only a global score on the final answer is given" (Rutherford & Watson, 1990:356).

In this study, final year-marks of the students in each subject were used as dependent variables and variance was studied for individual subjects rather than overall first year pass or fail, as it was felt that individual predictors might contribute differently for different courses (Rutherford & Watson, 1990:357).

The matric percentages obtained in English, Mathematics and Physical Science combined, explained 17-26% of the variance in the various first-year subjects. A multiple regression analysis was performed, using all of the mentioned predictors, and in this combination the variance accounted for ranges from 26-33% (Rutherford & Watson, 1990:358). The authors concluded that, the use of the selection battery in addition to the matric results, did increase the amount of variance in first-year grades explained, and have therefore increased the predictive validity of the selection process. They acknowledged however, that the particular tests used in the selection battery was not the best available and that first-order correlations of the results of these tests with end-of-the-year results proved insignificant. They admitted that new tests were being sought (Rutherford & Watson, 1990:358).

Smith (1992:5) proposed a model for the selection of civil engineering technician students at the Cape Technikon. In this study she randomly selected 145 students that were representative of the students in Civil Engineering at the Cape Technikon. She selected the following psychometric test battery:

- Senior Aptitude Test (SAT) - Verbal Reasoning
  - Three-Dimensional Spatial Perception
- Academic Aptitude Test (AAT) - Verbal Reasoning
  - Three-Dimensional Spatial Perception
- Gottschaldt Figures Test (for deductive reasoning)
- 16 Factor Personality Questionnaire (16PF)
- Self-Directed Research Questionnaire of Career Interests (SDS)
- Achievement Motivation Questionnaire (AMQ) by Pottas et al
- Job Satisfaction Index (JSI) by Brayfield and Rothe, adapted by Mauer
- A self-developed biographical questionnaire

Table 6.2: Correlations indicating factors with significant relation to the academic success of students in Civil Engineering (adapted from Smith, 1992:14)

Factor	Correlation r	p
Matric Mathematics	0.6183	0.0037**
Matric Physical Science	0.2731	0.0242*
Matric Mathematics and Physical Science	0.3607	0.0025**

Matric Mathematics, Physical Science and Languages	0.2746	0.0235*
Matric Mathematics, Physical Science and First Language	0.2748	0.233*
Matric Total Aggregate	0.5335	0.0044**
Three-Dimensional Spatial Perception (Raw score)	0.2259	0.0401*
AMQ (B)	0.3112	0.0039**
AMQ (AA)	0.3037	0.0326*
AMQ (BB)	0.2283	0.0404*
AMQ (PM)	0.3093	0.0050**
16 PF (E)	-0.2443	0.0432*
16 PF (I)	-0.2585	0.0336*
16 PF (L)	-0.3278	0.0210*
16 PF (N)	0.2703	0.0445*
16 PF (Q1)	-0.2621	0.0127*
16 PF (Q2)	0.6282	0.0016**

* *Significant on 5% level of meaning*

** *Significant on 1% level of meaning*

In Table 6.2, the factors that correlated significantly with the academic success of the sample are given. The matriculation results were converted to Swedish Formula scores by the method shown in Table 1 earlier in this chapter. In this study, Smith (1992:12) correlated the school results and psychometric data obtained with the final third-semester exam marks.

REMARK:

It is important to note that the sample for this study consisted of 91.7% male and 84.5 % White students (Smith, 1992:9). The sample is thus homogeneous and not representative of the national population.

From the results it is clear that, in academic terms, performance in matric Mathematics correlated the highest with the students' overall higher education performance (r=0.618; p=0.004). Furthermore, the total matric aggregate also correlated significantly (r=0.534; p=0.004) with the academic performance of the students.

As far as aptitudes are concerned, it seemed that only three-dimensional spatial perception, of those utilised, correlated significantly with the academic performance of the students (r=0.226; p=0.040). According to Smit (1996:234), the Spatial 3-D subtest on the Senior Aptitude Test measures primarily a "visualisation factor in which three-dimensional visualisation plays a role."

According to the results as shown in Table 6.2, four factors on Pottas, Erwee, Boshoff and Lessing's **Achievement Motivation Questionnaire** (1988) correlated statistically significantly with the academic performance of the students included in the study. Smit(1996:369) states that the AMQ was developed to, amongst others, measure the intensity of the motivation to achieve in adults.

> "Individuals who are achievement motivated strive to give of their best in whatever they undertake. They possess high personal standards of excellence and rely on their own abilities and skills to achieve success. In order to realise their goals, they persevere diligently, are action-orientated, and tend to manage time effectively and economically" (Smit, 1996:369).

In this particular study, the scores obtained for ***Time Consciousness*** (B) correlated the highest with the students' academic performance (r=0.311; p=0.003). Smit (1996:370) says that this subfactor indicates a person who plans and schedules his or her work. It furthermore seems that the score for ***Goal Orientation*** (AA) correlated significantly (r=0.304; p=0.032) with the academic performance of the students. People with high scores on this subtest tend to persevere despite obstacles and try to deal with matters as quickly as possible (Smit, 1996:370). In addition, the performance on the ***Personal Effectiveness*** subtest also correlated with the students' final exam marks (r=0.228; p=0.04). According to Smit (1996:370), high scorers on this factor usually set themselves very high standards of personal performance and have full confidence in their own abilities. The total ***Performance Motivation*** score also correlated significantly (r=0.309; p=0.005) with the academic performance of the respondents. Smit (1996:371) adds that it is advisable to take note of Factors AA and BB, as well as Factor PM, when using this questionnaire for selection purposes.

From Table 2 it can be seen that six of the factors of the **16 Factor Personality Questionnaire** (16PF) correlated significantly with the academic performance of the investigated group. The 16PF is claimed by Cattell (1961:2) to measure the total personality and was adapted by the HSRC for South African conditions and thus standardised (Smit, 1996:280).

The six factors or personality dimensions which correlated significantly with the academic performance of the students in Smith's study are briefly discussed below.

- **Factor E: Submissiveness versus Dominance**

  According to Smith (1992:14), Factor E showed an inverted correlation with the academic performance of students in her study (r=-0.244; p=0.043). This implies that the lower the score on Factor E, the better the academic performance. According to Smit (1996:292), a person scoring low on this factor might tend to be obedient, accommodating, considerate and conventional. An explanation for this correlation could be that these students were conscientious and therefore more successful.

- **Factor I: Tough-mindedness versus Tenderheartedness**

  The relation between the score obtained on Factor I by the students is again inverted to their academic performance (r=-0.259; p=0.033) (Smith, 1992:14). Smit (1996:294) states that "Persons who score low on this factor are realistic, accept responsibility, act on practical, logical grounds and are not artistic." In this case it seems as if the responsible, realistic students are the better performers in engineering courses.

- **Factor L: Trust versus Mistrust**

  The students' score on Factor L of the 16PF correlated significantly with their academic performance (r=-0.329; p=0.021) - again the relation is inverted. The lower the score on Factor L, the better the academic success. Smit (1996:296) states that people with a low score on Factor L tend to be relaxed and easygoing as well as trusting and accepting. It could be possible that these students coped better with the pressure of higher education studies.

- **Factor N: Artlessness versus Shrewdness**

  Smith (1992:14) reports a positive correlation between scores obtained on Factor N and the academic performance of engineering students (r=0.270; p=0.0445). Smit (1996:297) refers to people scoring high on this factor as being shrewd, emotionally detached and showing insight into self and others. This finding is not supported by previous findings, which revealed a negative relationship between this factor and success in education (Smit, 1996:297).

- **Factor Q1: Conservatism versus Radicalism**

  The relationship found by Smith (1992:14) between Factor Q1 and the academic performance of engineering students is inverted (r=-0.262; p=0.013). According to Smit (1996:298), persons measuring low on this factor are conservative and respect established ideas. This finding is again contrary to the expected, as research has shown that people with high scores on Factor Q1 manifest intellectual interests and tend to experiment.

- **Factor Q2: Group dependancy versus Self-sufficiency**

  Smith (1992:14) reports a significant correlation between Factor Q2 and the academic performance of the students taking part in her study (r=0.628; p=0.002). Smit (1996:298) sees people who score high on this factor as self-sufficient and resourceful; they furthermore tend to avoid social life, since they regard it as a waste of time. This may be the reason for their better academic performance.

Even though the sample used in this study was not representative of the general population and the results obtained therefore not generaliseable, some of the findings could serve as an indication of useful constructs that could be further investigated by this study.

Zaaiman (1998: 71) conducted a study in collaboration with the Free University of Amsterdam in which focus was placed on the selection of disadvantaged students for the ***University of the North Foundation Year*** (UNIFY). This foundation year is specifically aimed at preparing students for scientific and engineering courses by developing their Mathematical and Physical Science skills. According to Zaaiman (1998:74), 150 students are admitted to this course every year. Many more students apply and are evaluated. In 1993, 189 applicants were tested; from 1995 to 1997 the numbers varied from 600 to 700.

Zaaiman (1998:75) states that it was decided that the University of the North would use its own selection tests because of the uncertainty regarding the validity of the DET matric results for the prediction of future academic performance. These selection tests would be used to identify potentially successful students.

Critical performance areas were identified by relevant academic staff and test specifications were developed to match these. The test specifications were then used to develop the selection tests and procedures (Zaaiman, 1998:86). The areas that were assessed included the following:

- Mathematic skills (as evaluated by the UNIFY Mathematics Selection Test or UMST);
- Science skills (as evaluated by the UNIFY Science selection Test or USST);
- English proficiency (as evaluated by the UNIFY English selection Test or UEST);
- Basic arithmetic skills (as evaluated by the UNIFY Arithmetic Selection Test or UAST).

In this study, Zaaiman (1998:93) used the final average score obtained by students from the University of the North's Foundation Year as criterion for the calculation of the predictive validity coefficients. The year mark was used in cases where students were not allowed to write examination on the grounds of poor performance during the year.

Table 6.3 shows the predictive validity coefficients of the results of the selection tests for the final Foundation Year performance. (An initial English proficiency test was included in the table, since performance in this test was used to decide whether English language skills should be included in the test battery or not.)

Table 6.3: Predictive validity coefficients of the selection tests used at the University of the North's Foundation Year as calculated with the final academic performance as criterion (adapted from Zaaiman, 1998:93)

Predictor	1994 r	1995 r	1996 r
UNIFY Mathematical Selection Test	0.44**	0.39**	0.41**
UNIFY Science Selection Test	0.27**	0.38**	0.44**
Average Maths and Science	0.49**	0.49**	0.57**
UNIFY Arithmetic Selection Test	not available	not available	0.28**
English proficiency test	0.28**	0.27**	0.26**
UNIFY English Selection Test	not available	not available	0.30**

*** two-tailed significance less or equal to 0.01*

Further investigation showed that the best combination of predictors, as determined by a multiple regression analysis, were the Mathematics, Science and English Selection Tests. In 1996, the multiple R obtained by these predictors was 0.59, which was significant at $\propto = 0.01$ (Zaaiman, 1998:94).

The constructs identified in Zaaiman's study do in some instances correlate with those in Smith's study discussed earlier in this chapter, even though the type of assessment differed greatly. The role of all these constructs will be investigated thoroughly in this study.

6.7 Conclusion

From the literature reviewed it is clear that, although scholastic performance is shown as the single predictor with the highest predictive validity, a large portion of prospective students are placed at a severe disadvantage if this is used as sole selection criteria. As the majority of the applicants for the engineering technician courses at Technikon Pretoria come from the previously disadvantaged communities, and school performance is currently used as minimum admission requirement, the predictive validity of school performance as predictor of higher education performance will again be looked into.

The literature reviewed also revealed that the evaluation of learning potential does not significantly increase the prediction of academic success on higher education level. Furthermore, it is an impractical evaluation to carry out when large numbers of applicants have to be assessed. For these reasons, the evaluation of learning potential will not be included into the current study.

According to the literature studied, biographical factors have not been proved to be accurate predictors of academic performance at higher education level. However, it should be kept in mind, that it is difficult to subject these factors to statistical analysis, especially when the more conventional statistical techniques are used. It is thus possible that these results were not as accurate as they could have been. But, since this study is concerned with the development of a selection battery for first year engineering technology students, biographical variables such as age and gender cannot be included. It would be unconstitutional to deny applicants access to further studies on the grounds of age and/or gender.

As some success in the prediction of students' academic performance by means of ability tests has been documented, this study will concern itself with the identification and determination of the predictive validity of a test of this kind. Chapters 8 to 11 will deal with the different aspects of the research methodology followed in executing the study. Chapter 7 will look into the specific psychometric test used in the study.

Chapter 7

Potential Index Batteries

7.1 Introduction

Since some success in the prediction of students' academic performance by means of ability tests has been documented (see Chapter 6), it was decided to use a test of this kind in this study to determine the potential for success of prospective students in the engineering technician courses at Technikon Pretoria.

It was decided to make use of the Potential Index Batteries (PIB), developed by Erasmus and Minnaar for the purpose of this study. The aforementioned instrument had previously been purchased by the Department of Student Counselling at Technikon Pretoria and all the student counsellors had been trained by the developers and were familiar with the testing procedures. The Department of Student Counselling was satisfied that the initial research yielded sufficient evidence towards the instrument's culture fairness and compliance with all relevant regulations and legislation.

In this chapter, the nature, purpose and relevant sub-indices of the instrument will be discussed in detail. Reference will be made to studies carried out on the validity and reliability of the instrument.

7.2 The nature of the Potential Index Batteries (PIB)

According to Erasmus et al (1997:2), PIB can be described as "an instrument for the screening of the potential of literate, semi-literate, literate and academically advanced persons."

According to the test developers, the principle according to which the test was developed, was that **the ability to perform a simple task, indicates the potential to master a related, but more complicated task.**

The instrument consists of three principle sections, namely, a Comprehensive Structured Interviewing for Potential system (CSIP) as well as a visual or non-verbal (V-PIB) and a verbal (PIB) assessment section. Each of the latter sections consist of smaller units or indices - each measuring a specific competency. In total, the instrument can evaluate a person's potential regarding 65 different fields or competencies. The fields relevant to this study will be discussed in detail later in this chapter.

The Comprehensive Structured Interviewing for Potential system (CSIP) is used in the execution of a job analysis in order to compile a comprehensive job description. According to Erasmus (1997:46), the competencies relied on by CSIP were identified and generically defined from 1993 to 1995, in a joint project of the developers (Potential Index Associates) and the National Productivity Institute. By using the Job Specification Index it is possible to establish the most important predictors necessary for success in a particular job or task.

> "A systemized procedure leads to the ranking of the 65 competencies of the CSIP in their order of importance for performance in a particular job. It then proceeds to select a predetermined number of competencies - a minimum of 8 competencies are suggested here - which are regarded as the most vital ones for success in this particular position" (Erasmus, 1997:47).

According to the developers this procedure customises the assessment programme for each specific position, as candidates are only assessed on competencies vital for success in the particular position (Erasmus, 1997:47). It is furthermore essential that the persons involved in determining the relevant competencies should be familiar with the job (task) being analysed. The developers suggest that at least three, but preferably five or six people, should participate in the task. As soon as the most important competencies have been identified, the corresponding indices from the assessment sections (V-PIB and PIB) are included in the battery used for selection purposes.

7.2.1 Description of the relevant individual indices of the V-PIB and PIB

The following indices were identified as important for the prediction of the academic performance of first-year engineering technician students at Technikon Pretoria. The descriptions for the V-PIB indices are according to the User's Manual (Erasmus & Minnaar, 1995:1-6) and that of the PIB according to Erasmus et al (1997:17-20).

7.2.1.1 Visual Potential Index Batteries

- **Index 3 - Numerical Potential**
  This index comprises basic numerical calculations and is aimed at the respondent's mainly verbal potential to calculate. It is expected of the testee to reason along logical, mathematical lines, to apply reduction and induction, and it demands the capacity to add, subtract, divide and multiply.

- **Index 4 - Composition of Wholes**
  In this index, the testee has to synthesise small parts into an organised, integrated whole. It constitutes an important non-verbal component of potential. This capacity seems to involve both visual perception and logical reasoning. It furthermore demands a capacity for concrete

reasoning in terms of figures and forms and their logical place within a single structure. The testee has to be able to visualise the desired end-result in order to be able to perform the task.

- **Index 6 - Spatial Reasoning**
  In this index, the respondent is instructed to identify two identical illustrations amongst five almost similar ones. The capacity to perceive, analyse and synthesise is assessed, as well as the ability to select and categorise.

- **Index 7 - Perception**
  The respondent is expected to identify the one illustration that is different from four similar illustrations. This index assesses the potential to perceive detail as well as wholes in their specific, logical and sensible context.

7.2.1.2 Potential Index Batteries

- **Index 2 - Creativity**
  Creativity refers, in this context, to a person's need, willingness and ability to create new concepts and solutions to problems, rather than sticking to old and tried ways.

- **Index 3 - Reading Comprehension**
  The respondent is given five minutes in which to read a passage and is then asked questions on the contents of the passage. He or she is not allowed to return to the passage once the questions are being answered. This index assesses the testee's competency to read and understand clearly what the reading matter conveys.

- **Index 5 - Mental Alertness**
  This index requires the respondent to identify the sole deviation from five seemingly related objects. This ability is, according to the developers,

associated with the ability to classify objects correctly. Schaap (1997:70) refers to the fact that this index is also associated with the general mental ability (or G-factor) of the testee.

- **Index 12 - Vocabulary**
  The respondent's basic English vocabulary is assessed in this index. He or she is expected to select, from five possible answers, a synonym for a specific word given.

7.3 Validity and Reliability of PIB

The developers of PIB take great pride in the fact that they have made the instrument available for scrutinising by both its users and experts from the academic sphere. Studies for determining the validity and reliability of PIB have been undertaken by various persons from various institutions.

Correlations between PIB indices and other standardised psychometric tests were found to be high. In this respect, Schaap (1997:72) found correlations between PIB and the 16 Personality Field Questionnaire as high as 0.70 ($p = 0.01$).

In a run-up to the study done here, correlations were calculated between the PIB indices and the Academic Aptitude Test (AAT). An accidental sample of 500 psychometric records of prospective students was used for this study. According to Dane (1990:302), an accidental sample is based on availability or ease of inclusion. In the case of this study, these records were available at Technikon Pretoria, due to an organisational search for suitable psychometric instruments for use in selection.

It was found that Index 4 of the V-PIB (Composition of Wholes) correlated significantly with Subtest 1 of the AAT (Non-Verbal Reasoning) and with Subtest 9 (Spatial Perception). In the case of Subtest 1, a correlation of 0.49 ($p = 0.01$) was found and in the instance of Subtest 9, a correlation of 0.48 ($p = 0.01$) was found. Index 6 (Spatial Reasoning) of V-PIB correlated at 0.39 ($p = 0.05$) with Subtest 9 (Spatial Perception) of the AAT. PIB's Index 5 (Mental Alertness) correlated with the AAT's Subtest 1 ($r = 0.58$; $p = 0.01$), Subtest 2 ($r = 0.71$; $p = 0.01$) and Subtest 9 ($r = 0.46$; $p = 0.01$).

The conclusion arrived at from the above results was that the instrument did show construct validity in previous studies and could therefore be used in this study. The calculation of the situation-specific validity for the current study will be discussed in a later chapter.

In a study executed by Schaap (1997:4), reliability coefficients of as high as 0.93 were found for Index 5 (Mental Alertness) of the PIB. Pilot studies undertaken by Technikon Pretoria rendered reliability coefficients for indices ranging from 0.69 to 0.96. The determinating of situation-specific reliability coefficients, relevant to the indices used in this study, will be discussed in Chapter 12.

7.4. Conclusion

Two factors influenced the decision to use the Potential Index Batteries in the development of a selection battery for engineering technician students at Technikon Pretoria. Firstly, the availability of the instrument was important, but even more important was the fact that the instrument was developed in South Africa and regarded as culture friendly. This made the instrument acceptable to student bodies and the Technikon Management, and also acceptable from a professional and legal point of view.

In Chapter 8 the methodology used in the research process will be discussed.

Chapter 8
Research Methodology

8.1 Introduction

As mentioned in Chapter 1, the objective of this study is the development and validation of a selection battery for first-year engineering technician students at Technikon Pretoria.

Possible predictor variables have to be identified, and, in order to determine the predictive validity of these predictor variables, a correlation has to be drawn between predictor scores and criterion scores. The predictive validity thus calculated will be an indication of the accuracy with which the predictors predict criterion performance.

In the following four chapters, attention will be given to the discussion of the research methodology followed in this study. Aspects which will be discussed in detail in this chapter will be the research design and the data collection techniques used in the study. In Chapter 9, the population from which the sample for this study was drawn will be illustrated, together with the sampling method and the sample itself. Chapter 10 will deal with the statistical techniques used in the data analysis process. Chapter 11 will explain the determination of the predictor variables.

8.2 Research Design

According to Guy, Edgley, Arafat and Allen (1987:92), a research design can be defined as "the plan of procedures for data collection and analysis that are undertaken to evaluate a particular theoretical perspective." McMillan and Schumacher (1993:157) add that research design refers to a plan for selecting subjects, research sites and data collection procedures to answer research questions. They emphasise that the goal of a sound research design is to ensure results that are credible.

The design selected for this study is a non-experimental, correlational design (Smit, 1983:70; McMillan & Schumacher, 1993:270). According to Smit (1983:70), this research technique is considered the best controlled and most accurate of all non-experimental designs. This design specifically deals with the determination of the extent and intensity of covariation or interdependency between variables in natural circumstances (Smit, 1983:70).

McMillan and Schumacher (1993:270) distinguish between two types of correlational research, namely, simple ***correlational studies*** and ***prediction studies***. They define a simple correlation study as a research study where the researcher obtains scores "from two variables for each subject and then uses the pair of scores to calculate a correlation coefficient" (McMillan & Schumacher (1993:270). In a prediction study, one variable (**predictor variable**) is used to predict performance on a second variable (**criterion variable**) and the predictor variable is usually measured before the criterion variable (McMillan & Schumacher, 1993:271). A ***positive correlation*** means that high values on one variable are associated with high values on a second variable, whereas a ***negative correlation*** is found when a high value on one variable is associated with a low value on a second variable (McMillan & Schumacher, 1993:35).

Smit (1983:70) states that correlative techniques deal with variables already existing within the subjects' repertoire and **not** with variables generated experimentally. He continues that correlational research differs from experimental research in so far as that in correlational research no manipulation of an independent variable occurs, and therefore no distinction can be made between an experimental and a control condition.

Smit (1983:74) regards psychometric test construction and research as the most general use of the correlational technique as a research method.

8.3 Data collection

According to Guy et al (1987:112), data collection refers to the tool or vehicle through which measurement is actualised. They add that even though the research design does not dictate the data collection procedures, certain data collection techniques usually accompany certain designs.

Smit (1983:145) as well as McMillan and Schumacher (1993:40) distinguish between quantitative and qualitative techniques of data collection. McMillan and Schumacher (1993:14) state that, on one level, the terms refer to distinctions made about the nature of knowledge and on another level they refer to the research method - that is, how data are collected and analysed - as well as the type of generalisations derived from the data. "**Quantitative research** presents statistical results represented with numbers; **qualitative research** presents facts in narration of words" (McMillan & Schumacher, 1993:14).

Quantitative techniques are used with experimental, descriptive and correlational designs (McMillan & Schumacher, 1993:41). As the design selected for this study is correlational, the data collection techniques will subsequently be quantitative.

McMillan and Schumacher (1993:167) name the following as common methods of gathering quantitative information:

- ⌘ Questionnaires
- ⌘ Standardised interviews
- ⌘ Tests
- ⌘ Standardised observations
- ⌘ Inventories
- ⌘ Rating scales

In the current study, two methods of data collection were used. The first involved a questionnaire that was used to identify the variables that were to be included in the prediction study. McMillan and Schumacher (1993:42) refer to questionnaires as a technique that encompasses a variety of instruments in which the subject reacts to written questions.

The questionnaire used in this study is a standardised questionnaire that forms part of the Potential Index Batteries discussed in Chapter 7. It was decided to analyse the requirements the Engineering Technology training programme expected the incoming students to meet by focussing on the actual performance of students in the programme. As a starting point in this process, the Comprehensive Structured Interviewing for Potential system (CSIP) was completed by a group of lecturing staff from the Engineering Faculty at Technikon Pretoria. This group of respondents consisted of 12 members of the lecturing staff, who are all directly involved with the training of Engineering Technology students.

The questionnaires asked the respondents to select, from a list of 65 competencies the eight which are indispensable for the success of an engineering technology student. Subsequently the respondents were asked to rate these eight competencies in order of significance. These responses were then computer- processed and the results indicated to the researcher which predictive variables were thought important for the success of students by academic staff concerned with the training of engineering technology students. In this manner, the competencies to be included in the potential assessment battery were selected.

For the assessment of these competencies, psychometric testing was selected (as discussed in Chapter 6). This was the second data-collecting method to be implemented by the researcher. According to McMillan and Schumacher (1993:42), "testing" refers to the use of tests scores as data. "This technique involves subject response to either written or oral questions to measure knowledge, ability, aptitude, or some other trait. A numerical value is obtained as a result of each subject's answers to a standard set of questions" (McMillan & Schumacher, 1993:42).

According to Smit (1983:169), methods can be regarded as objective when different scorers give the same numerical values to the testees when following a prescribed set of rules. Psychometric testing and its characteristics were discussed in detail in Chapter 4 and only a summary of the uses of objective psychological tests - as stated by Smit (1983:169) - will be given here.

8.3.1 Uses of psychological tests in research

a. Psychological tests can be used for selection purposes. In other words, a person can, on the grounds of his or her performance on a test, be included or excluded from a study, grouped in the experimental or control group, etc.

b. Psychological tests can be used to determine the effect of the experimental/independent variable.

The psychometric test used in this study was the Potential Index Batteries (PIB) of Erasmus and Minnaar (1995). (This test was discussed comprehensively in Chapter 7.) Students who applied for either the National Diploma In Civil Engineering Technology or the National Diploma in Mechanical Engineering Technology were evaluated by means of the PIB. These assessment sessions were run by staff members of the Department of Student Counselling at Technikon Pretoria. All of these staff members are registered with the Health Professions Council of South Africa as either psychometrists or counselling psychologists. They have also been trained in the administering of the PIB by the developers themselves. The psychometric results of the students who had been admitted to these courses were included in the study, since academic records of these students, which could be used as criteria variables in the study, would be available. These academic results consisted of the final first-semester mark obtained by students in all their required subjects.

The final matric symbols obtained in Mathematics and Natural Science were obtained from the Academic Administration Division at Technikon Pretoria. The matric symbols were quantified according to the Swedish formula discussed in Chapter 6, and added to the equation as predictor variable.

8.4 Data analysis

According to Guy et al (1987:308), the intent of the data analysis phase in research is to examine the body of data for the hypothesised relationships. It is in this stage of the research that answers to the research questions are gained.

Guy et al (1987:309), as well as McMillan and Schumacher (1993:191,) states that quantitative research findings, as in the case of this study, are usually presented in statistical form. The statistical techniques used in the data analysis phase of this study will be described in Chapter 10.

8.5 Conclusion

In providing an answer to the research objective stated in Chapter 1, a research design was implemented which would best suit the type of investigation under way in this study. A nonexperimental, correlational design was selected, as this research technique is considered the best controlled and most accurate of all nonexperimental designs (Smit, 1983:70). Since a quantitative technique was selected for data gathering, the necessity for a statistical method in the data analysing process was obvious. The results of the data analysis will be discussed in detail in Chapter 12.

Chapter 9

The sample

9.1 Introduction

One of the first steps when designing quantitative research is to choose the subjects who will participate in the study (McMillan & Schumacher, 1993:159). Theorists seem to agree that this sampling process constitutes one of the most important steps in a research project, as it will determine the generalisability of the research findings (Smit, 1983:178; Guy et al, 1987:174; McMillan & Schumacher 1993:160). In this chapter the population from which the sample for this study was drawn will be illustrated, together with the sampling method and the sample itself.

9.2 The population

Dane (1990:289) describes a population as "all possible units or elements that can be included" in a study or research project. In the current study, the population comprised all mechanical, electrical and civil engineering technology students at Technikon Pretoria from, 1997 to 1999, of whom psychometric records were available.

9.3 The sample

Guy et al (1987:174) defines sampling as a method of selecting some part of a group to represent the total. The objective of sampling is the estimation of population values from the information contained in the sample. Theorists such as Smit (1983:182), Guy et al (1987:184) and McMillan and Schumacher (1993:160) distinguish between two major sampling procedures: probability and non-probability samples. In the case of probability samples, the researcher can specify for each unit in the population the likelihood that it will be included in the sample. With non-probability samples, there are units of the population that have no chance of being included in the study.

9.3.1 Factors to be taken into consideration when sampling

Smit (1983: 179) regards the following factors as crucial when selecting a sampling procedure:

- The statistical definition of the research question as well as the type of information needed.
- The definition of the population in order that elements belonging thereto can be easily identified. This enhances the validity of the sample.
- Similar surveys ought to be studied to ensure that information does not already exist.
- After the nature of the data required has been identified the best process for sampling should be selected.

9.3.2 The sample size

According to McMillan and Schumacher (1993:163), the researcher must determine the size of a sample that will provide sufficient data to answer the research question. These authors, as well as Guy et al (1987:196), mention that the general rule for determining sample size is to use the largest sample possible. McMillan and Schumacher (1993:165) continue that in the case of correlational research - as in this study - a minimum of thirty subjects should be included. If the researcher expects to find small differences or slight relationships, they feel that it is desirable to have as large a sample as possible. A large sample is also needed when a study concerns itself with a variety of dependent and/or independent variables. In the case of this study, it will thus be preferable to have the largest sample that is available.

9.3.3 The sample for this study

For the purpose of this study, a non-probability convenience sample was selected. Smit (1983:195) states that this method implies that the researcher uses whatever subjects are available for the study. Guy et al (1987:189) states that the major deficit of this method is that there is no formal procedure for generalising from sample to population, since sampling error cannot be determined. On the other hand, using a convenience sample could save time and money - thus, what is lost in accuracy would be gained in efficiency. Smit (1983:195) concludes that this sampling method is totally satisfactory if the researcher does not plan to generalise findings to the broad population.

The sample for this study consisted of a total of 732 Engineering Technology students at Technikon Pretoria. From these, 512 were Civil Engineering Technology students and the remaining 220 were Mechanical Engineering Technology students. These subjects were the total number of students from these two academic departments, enrolled from 1997 to 1999, of whom both psychometric and academic data were available.

As depicted in Figure 9.1, the sample consisted of 14.75% female and 85.25% male respondents. Figure 9.2 shows the frequency distribution of the home languages spoken by the respondents. The respondents indicated twelve different languages (frequency shown in brackets):

English	(18.72%)
Afrikaans	(29.64%)
Zulu	(4.92%)
Ndebele	(2.73%)
Tsonga	(4.37%)
Northern Sotho	(14.89%)
Venda	(3.83%)
Tswana	(9.29%)
Xhosa	(2.32%)
Southern Sotho	(4.23%)
Portugese	(0.96%)
Swazi	(0.68%)
Other	(0.96%)

Of the respondents, 5.24% indicated qualifications higher than a National Senior Certificate. The average age of the group was 18.08 years with a standard deviation of 2.04.

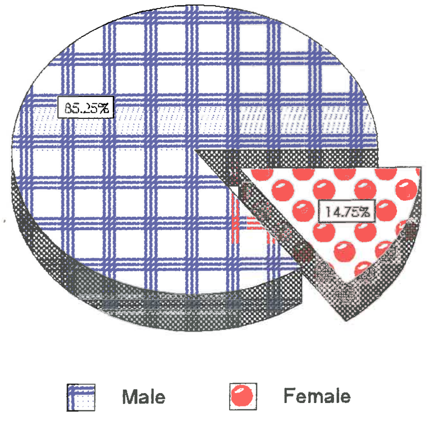


Figure 9.1: Frequency distribution of the gender of the sample used in the study

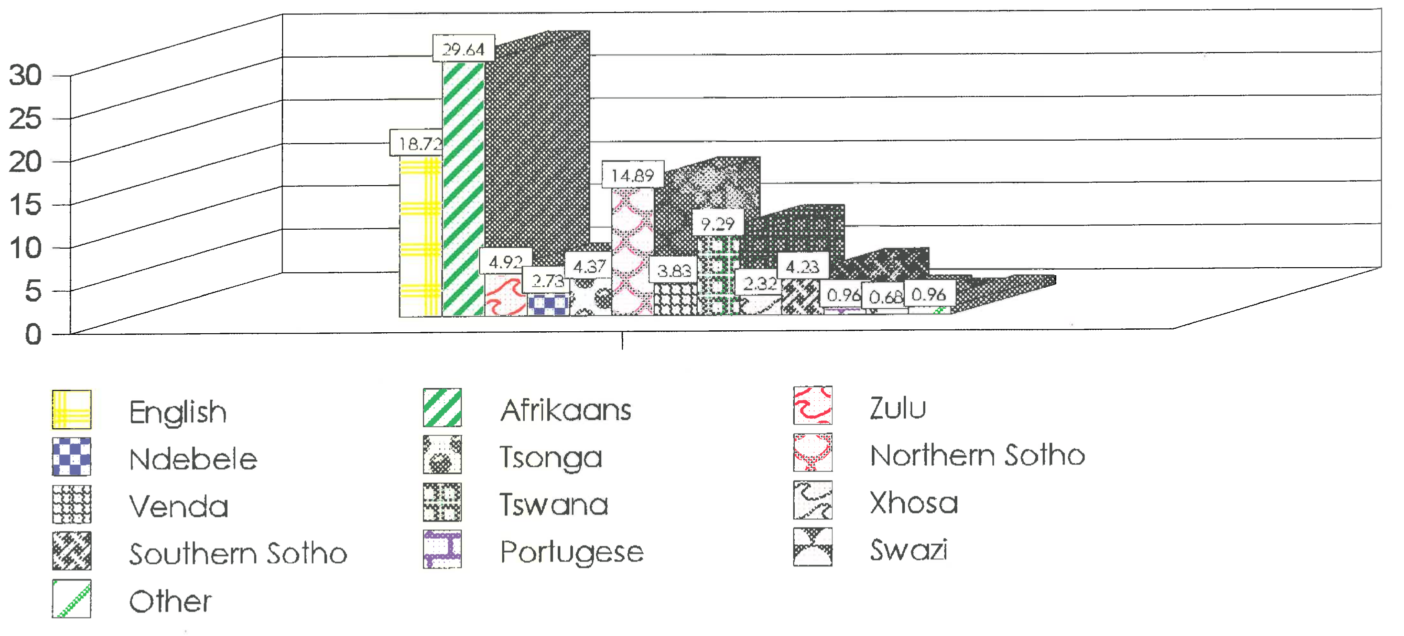


Figure 9.2: Frequency distribution of home languages spoken by respondents

9.4. Conclusion

As De la Rey (1976:159) states, sampling is decisive in the research process. It largely determines the success of a scientific study. In Chapter 9, the sampling method used in this study was described and motivated, and the sample itself was described. As a nonprobability sample was used, it was not possible to determine sampling error (Guy et al, 1987:196) and therefore as many subjects as were available were included in the sample. Chapter 10 will focus on the statistical techniques used in the data analysis process.

Chapter 10
Statistical techniques

10.1 Introduction

Guy et al (1987:309), as well as McMillan and Schumacher (1993:191), states that quantitative research findings, as in the case of this study, are usually presented in statistical form. Two major approaches comprise the subject matter of statistics, namely, **descriptive** and **inferential** (Howell, 1992:4).

McMillan and Schumacher (1993: 192) say that descriptive statistics transform a set of numbers and/or observations into indices which describe or characterise the data. Inferential statistics are used to make inferences or predictions about the similarity of the sample to the population from which the sample was drawn. The two approaches are interlinked and McMillan and Shumacher (1993:192) conclude: "...a researcher would first take a sample from a population, use descriptive statistics to describe the sample, and then use inferential statistics to estimate the true value of the test score for the population."

According to Dane (1990:243), statistical analyses used for predictive research are generally based on correlation, that is, the measure of the extent to which two or more variables are related. These methods of analysis, inter alia, include correlation coefficients, simple regression analysis and multiple regression analysis.

The descriptive and inferential techniques used in this study will subsequently be discussed.

10.2 Descriptive statistics

As mentioned in the introduction, descriptive statistics transform a set of numbers and/or observations into indices which describe or characterise the data. For the purpose of this study, measures of central tendency, measures of variability and frequency distribution are important.

10.2.1 Measures of central tendency

Bless and Kathuria (1993:35) indicate that measures of central tendency are values usually used to give a general description of a bulk of data. Theorists distinguish between three principle measures of central tendency, namely, mode, median and mean (Bless & Kathuria,1993:35; Howell, 1992:31-32; Popham & Sirotnik, 1992:14-16).

10.2.1.1 The mode

The mode is the score that occurs most frequently in a distribution (Howell, 1992:31; Popham & Sirotnik, 1992:16; Bless & Kathuria, 1993:35). According to Popham & Sirotnik (1992:16), the mode is the index of central tendency least used. The mode has, according to Howell (1992:32), the advantage of not being influenced by extreme scores.

10.2.1.2 The median

The median is the midpoint in a set of ranked scores (Howell, 1992:31; Popham & Sirotnik, 1992:15). Howell (1992:32) states that the median shares it major advantage with the mode in that it is unaffected by extreme scores on either end of the distribution. Guy et al (1987:322) add that the median becomes an especially important measure when there are a few extreme scores in a distribution.

10.2.1.3 The mean

The most often used measure of tendency is the mean, which is actually the arithmetic average of a set of data (Howell, 1992:32; Popham & Sirotnik, 1992:14). According to Howell (1992:33), certain disadvantages are associated with the mean. It is influenced by extreme scores and its value may not actually exist in the data. In its favour it can be said that it is unbiased and efficient and can be manipulated algebraically (Howell, 1992:33). Bless and Kathuria (1993:46) add that since the mean makes use of every score in the distribution, it is the most accurate measure of central tendency.

10.2.2 Measures of variability

Guy et al (1987:322) state that while measures of central tendency are concerned with how the scores in a distribution are grouped together, measures of variability are concerned with how the scores in a distribution are spread apart. McMillan and Schumacher (1993:205) add that measures of variability indicate how much scores, on the average, differ from the mean.

10.2.2.1 Variance

Bless and Kathuria (1993:63) as well as Guy et al (1987:325), define variance as the mean of the squared deviations from the distribution's mean. In formula form the variance can be expressed as:

$$s^2 = \frac{\sum (X - \bar{X})^2}{N}$$

10.2.2.2 Standard deviation

McMillan and Schumacher (1993:206) state that the standard deviation is a numerical index that indicates the average variability of the scores; it thus gives information about the distance, on the average, of the scores from the mean. Bless and Kathuria (1993:64) define the standard deviation as the positive square root of the variance. In formula form the standard deviation can be expressed as:

$$s = \sqrt{\frac{\sum (X - \bar{x})^2}{N}}$$

10.2.3 Frequency distribution

In a frequency distribution the data are listed or grouped in order to assess how often a particular score or result occurs (Bless & Kathuria, 1993:vi). Usually the frequency distribution of a set of data is presented graphically.

10.3 Inferential statistics

According to Bless and Kathuria (1993:77), inferential statistics refer to techniques for making statements on the basis of partial information. They add that the aim of inferential statistics is to provide information that is not directly included in the data. For this study, the determination of correlation coefficients and the performing of multiple regression analysis are important and they will thus be discussed.

10.3.1 Correlation coefficients

According to Popham and Sirotnik (1992:65), the term "correlation", in statistical analysis, always refers to a quantifiable relationship between two variables. McMillan and Schumacher (1993:215) state that the most common correlation technique is the Pearson product-moment coefficient. The product-moment correlation is used when both of the variables use continuous scales.

10.3.2 Multiple regression analysis

Hair, Anderson, Totham and Black (1992:7) describe multiple regression as the method of analysis that is appropriate when the research problem involves a single metric dependent variable presumed to be related to one or more metric independent variables. The objective of multiple regression analysis is to use the independent variable of which the values are known to predict the single dependent value the researcher wishes to know. The result is a variate, a linear combination of the independent variables that best predicts the dependent variable (Hair et al, 1992:25).

Popham and Serotnik (1992:104) argue that the technique of regression enables investigators to make predictions regarding a person's performance on one variable, given that person's performance on another variable. In multiple regression then, two or more predictor variables, related to the criterion variable, are incorporated into a more complex prediction scheme. The authors motivate their argument as follows:

> "...by adding a second variable related to the criterion variable, it is possible to reduce the standard error of estimate that would be present in a single predictor scheme. The better additional predictor usually is one that is at the same time related to the criterion variable and not too strongly related to the predictor variable already used." (Popham & Serotnik, 1992:105.)

McMillan and Schumacher (1993:273) state that the combined effect of the independent variables, in terms of the predictive power, to the dependent variable is presented by **R**, the coefficient of multiple correlation. The coefficient of multiple correlation can be thought of as simple correlation of all the independent variables together with the dependent variable. According to Hair et al (1992:64,) this coefficient reflects only the degree of association between two variables. **R square** (R^2) is the correlation coefficient squared, also referred to as the coefficient of determination. This value indicates the percentage of total variation in the dependent variable explained by the independent variable (Hair et al, 1992:64).

In most instances of multiple regression the researcher has various possible independent variables to select for inclusion in the regression equation. In the process of identifying the best regression model for the situation, the researcher can follow one of two general approaches, namely, a sequential search process or a combinatorial method (Hair et al, 1992:56).

Two general types of sequential processes are backward stepwise regression and forward stepwise regression (Hair et al, 1992:57).

Backward stepwise regression involves computing a regression equation with all the predictor variables and then going back and deleting independent variables that do not contribute significantly.

Forward stepwise regression also allows the researcher to examine the contribution of each predictor variable to the regression model independently. Each variable is considered for inclusion in the model prior to developing the equation.

According to Hair et al (1992:58), the combinatorial approach is primarily a generalised search process across all possible combinations of independent variables until such time as the best-fitting set of variables is identified.

As it was extremely important to know the contribution of each independent predictor variable to the regression model, the forward stepwise process were selected for this study. By being able to identify each individual predictor's contribution to the regression model, it should be possible to select the predictors with the highest validity to include in the final assessment battery, which is the ultimate goal of this study.

10.3.2.1 The standard error of estimate

The **standard error of estimate** functions as a measure of accuracy of predictions. Hair et al (1992:64) describe the standard error of estimation as "...the square root of the sum of the squared errors divided by the degrees of freedom. It represents an estimate of the standard deviation of the actual dependent values around the regression line; that is, it is a measure of variation around the regression line."

10.4 Levels of significance

McMillan and Schumacher (1992:341) state that since the basis of inferential statistics is the probability of estimation, the acception or rejection of the null hypothesis is also related to probability of chance. The **level of significance** indicates what the chance is that the researcher is wrong in rejecting the null hypothesis. In this study a level of significance of 5% was accepted. That implicates that assumptions made should be proven correct in 95 out of 100 instances. According to De la Rey (1976:172) researchers, ultimately aim at significance on the 1% and especially the 0.1% level, but all three of the levels mentioned are accepted in scientific research.

10.5 Steps taken in statistical analysis

The statistical analysis in this study involved the following steps:

✍ The mean and standard deviation for all predictive and criteria variables were calculated. A frequency distribution for each variable was drawn. Attention was given to the median, as well as the skewness and kurtosis of each distribution.

- ✍ A multiple regression analysis on the two sets of data was performed.

- ✍ The results of the senior students were analysed, again using multiple regression, in the cross validation study.

- ✍ The level of significance of the results was calculated and discussed.

- ✍ A battery for the selection of engineering technician students at Technikon Pretoria was then suggested.

10.6 Conclusion

Quantitative research always implies statistical analysis. In this chapter the statistical techniques that played an important role in the data processing procedure followed in this study were discussed. Attention was specifically given to multiple regression analysis, as this technique is important in predictive studies. The results obtained in the analysis will be discussed in Chapter 12.

Chapter 11
Determining the predictor variables

11.1 Introduction

From the literature overview reported in Chapter 6, it is clear that a large number of possible predictors for academic performance exist. All of them cannot be tested in the practical process of selection and not all will contribute to the fairness and predictive validity of the selection battery. As the main aim of this study is to find the best combination of practically measurable predictors to include in a selection battery for Engineering Technology students at Technikon Pretoria, a specific instrument was used to identify some of these predictors. The results obtained by means of this instrument were combined with some of the variables found to be valid in similar studies. The determination of the predictor variables will be discussed in detail in this chapter.

11.2 The Comprehensive Structured Interviewing for Potential System

As described in Chapter 8, the actual performance of students in the programme was analysed in order to determine the competencies important for success in an Engineering Technology training course. As a starting point in this process the Comprehensive Structured Interviewing for Potential system (CSIP) was completed by 12 lecturers of the Faculty of Engineering, who are all directly involved in the training of Engineering Technology students.

The questionnaires asked of the respondents to select the eight competencies from a list of 65 which are indispensable for the success of an Engineering Technology student. Subsequently, the respondents were asked to rate these eight competencies in order of significance. One of the essential areas in this process was to ensure that the competencies referred to requirements set for the incoming student and not for the required student level at the end of the training programme.

These responses were then computer processed and the results indicated to the researcher which predictive variables were thought to be important for the success of students by academic staff concerned with the training of engineering technician students. The competencies indicated by the respondents were:

- Creativity
- Reading Comprehension
- Mental Alertness
- Vocabulary
- Numerical Ability
- Composition of Wholes
- Spatial Reasoning
- Perception

Chapter 7 gives a detailed description of each of these competencies, as defined by the programme developers.

11.3 Academic criteria

At present the admission requirement for technikon education is a Senior Certificate. In the case of Engineering Technology courses at Technikon Pretoria the Senior Certificate should include Mathematics and Physical Science as subjects. A prospective student should furthermore obtain at least a C symbol on the Standard Grade or an E symbol on the Higher Grade to be eligible for admission.

From the literature reviewed it is clear that, although scholastic performance is shown as the single predictor with the highest predictive validity, a large portion of prospective students are at a severe disadvantage if this is used as sole selection criteria. As the majority of the applicants for the Engineering Technology courses at Technikon Pretoria come from the previously disadvantaged communities, and school performance is currently used as minimum admission requirement, the predictive validity of school performance as predictor of higher education academic performance will again be looked at.

11.4. Conclusion

The competencies indicated by academic staff involved with the training of Engineering Technology students at Technikon Pretoria were hypothesised to be indicative of a potentially successful student. After the identification of these predictor variables the assessment battery to be used in this study was compiled. This was then included in a comprehensive set of data regarding each applicant, together with the required school performance. Chapter 12 will describe the results obtained after a statistical analysis on the total set of data had been performed.

Chapter 12
Research findings

12.1 Introduction

The aim with the current study was to establish a potential assessment battery for the Engineering Technology courses at Technikon Pretoria that would be accurate, effective, fair and efficient. The Technikon Pretoria Potential Assessment Battery for Engineering Technology students was evaluated on the grounds of empirical research data collected in the period 1996 to 1999. The determination of the predictive validity of school performance (as indicated by symbols obtained), formed part of this study, as this still forms an integral part of the selection procedure at Technikon Pretoria.

In this chapter, the results of the study will be reported, as they relate to the objectives stated in Chapter 1. The chapter contains the descriptive statistics of the various predictive variables, the reliability of the indices used for the potential assessment and the predictive validity of the variables as determined by multiple regression analysis.

12.2 Descriptive statistics of predictor variables

Table 12.1 contains a summary of the descriptive statistics of the results obtained by the respondents on the indices of the Potential Index Batteries included in the assessment of prospective Engineering Technology students at Technikon Pretoria.

Table 12.1: Descriptive statistics of results obtained by the respondents on the PIB indices used in the study (maximum score possible indicated in brackets).

Index	Mean	Median	Standard Deviation
Creativity (100)	46.705	47.500	4.300
Reading Comprehension (100)	65.873	68.000	14.044
Mental Alertness (100)	83.139	84.000	6.919
Vocabulary (100)	78.759	86.000	17.710
Numerical Ability (10)	5.771	5.000	1.629
Composition of Wholes (10)	6.519	7.000	2.550
Spatial Reasoning (10)	5.089	5.000	1.531
Perception (10)	6.392	6.000	2.124
Mathematics (5)	1.983	2.000	1.174
Physical Science (5)	1.903	2.000	1.046

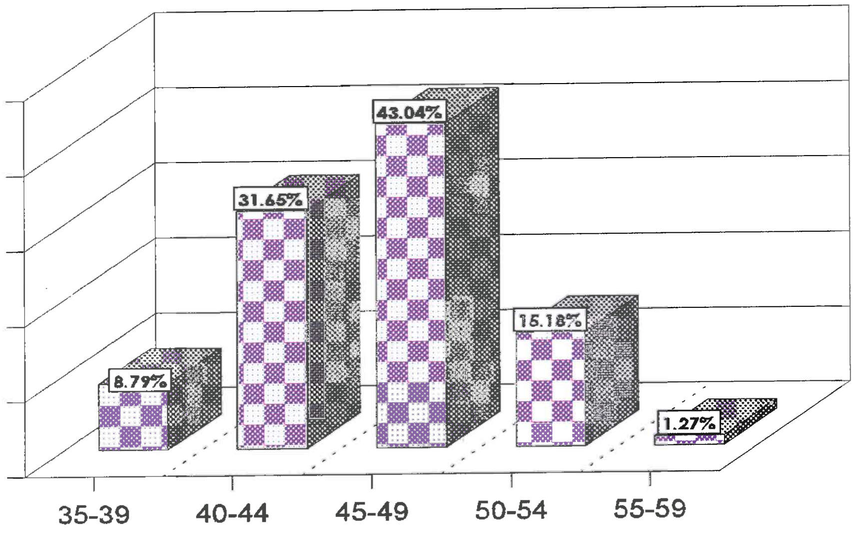
When interpreting the descriptive statistics, it is important to realise that these statistics describe the scores obtained by a preselected group. Only the scores of those students who were finally admitted to either the Civil Engineering Technology or the Mechanical Engineering Technology course were included in the analysis. It is inevitable that the distributions of scores will not be normal in all instances. The researcher took cognizance of the fact that this might influence the validity of the prediction models.

Table 11.1 indicates that in almost all of the indices used the median is higher than the mean of scores obtained (with the exception of the Numerical Ability, Spatial Reasoning and Perception indices). The majority of respondents can thus be said to have performed better than the mean. This phenomenon could most likely be explained by the fact that most of the applicants who performed below the average on more than half of the indices were not admitted to the relevant courses and their scores are thus not represented by these statistics. There were, however, applicants who obtained low scores on one or two indices only, and who were admitted to the Engineering Technology courses. Their scores are included in this set of descriptive statistics. As these scores were sometimes extremely low, they influenced the mean calculated for that particular set of scores.

As explained in Chapter 10, the standard deviation provides information about the distance, on the average, of the scores from the mean. It can thus be deduced from the information given in Table 12.1 that a wider range of scores was obtained on the Reading Comprehension and Vocabulary indices. This could be explained by the fact that students who performed lower than average on these two indices only were not rejected, but were conditionally admitted (the condition being that they receive formal study guidance and English literacy intervention).

The frequency distribution of the results obtained by the respondents included in the sample are presented graphically in Figures 12.1 to 12.8.

Figure 12.1: Frequency distribution of results obtained on the Creativity index of the PIB.



From Figure 12.1 it can be seen that the highest frequency of results obtained by the students on the Creativity index fell into the 45-49 range (43.04%). It can furthermore be seen that the range in which the respondents scored was limited between 35 and 59. In other words, the students obtained neither extremely low scores nor extremely high scores on this index. This could also be seen from Table 12.1 as the difference between the mean and median were relatively small, as was the standard deviation of the scores obtained on this index.

Figure 12.2: Frequency distribution of results obtained on the Reading Comprehension index of the PIB.

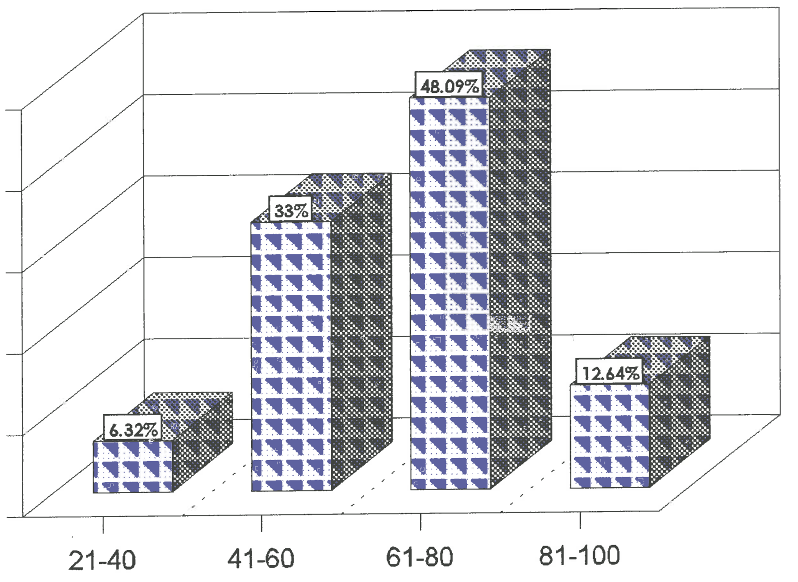


Figure 12.3: Frequency distribution of results obtained on the Mental Alertness index of the PIB.

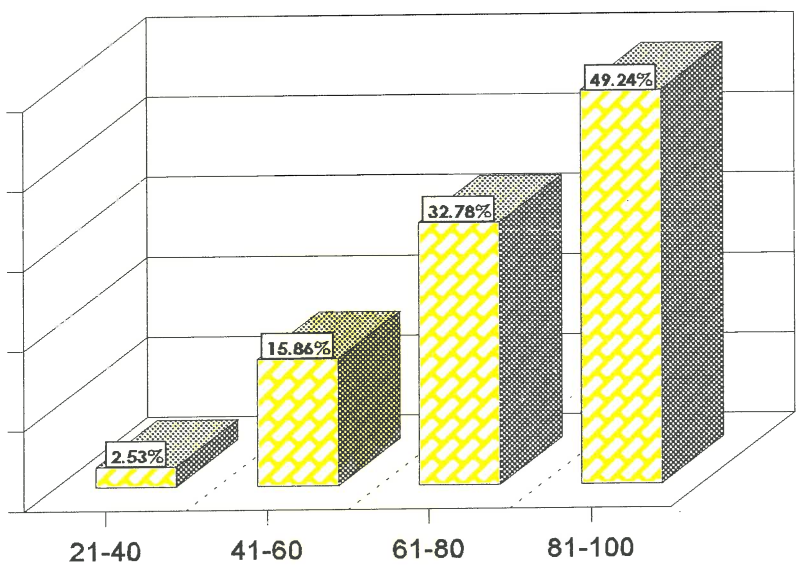


Figure 12.2 shows that the majority of respondents (48.09%) obtained results in the 61-80 range. It is also clear that a fairly large proportion of the respondents

obtained scores that were considerably lower than the mean, as 39.23% of the respondents obtained a score lower than or equal to 60 ($\bar{x}$ = 65.87). The range of results that were obtained varied between 21 and 100, which confirms the deduction made from Table 12.1 concerning the higher standard deviation. The fact that applicants were admitted to the various Engineering Technology courses despite performing much lower than average on this index can be seen clearly in this graph.

Figure 12.3 indicates the frequency distribution of the results obtained on the Mental Alertness index by the students admitted to the Engineering Technology courses. It is clear from the graph that the performance of the greatest majority of students' (49.24%) fell into the 81-100 range. Due to research done in other sectors, which indicated that this specific index has a high predictive validity for work performance, the counsellors who assessed the performance were much less lenient regarding an applicant's performance on this index. The result of this is quite clear from Figure 12.3 - the distribution indicates that very few applicants who did not perform at least average on this index were admitted to the various Engineering Technology courses.

Figure 12.4 shows that the majority of students performed better than average on the Vocabulary index of the PIB. Once again this shows that, in the assessment of the results, students who performed below average on this index were still admitted to the Engineering Technology courses. This is also confirmed by the standard deviation of 17.71 calculated for this set of data.

Figure 12.4: Frequency distribution of results obtained on the Vocabulary index of the PIB.

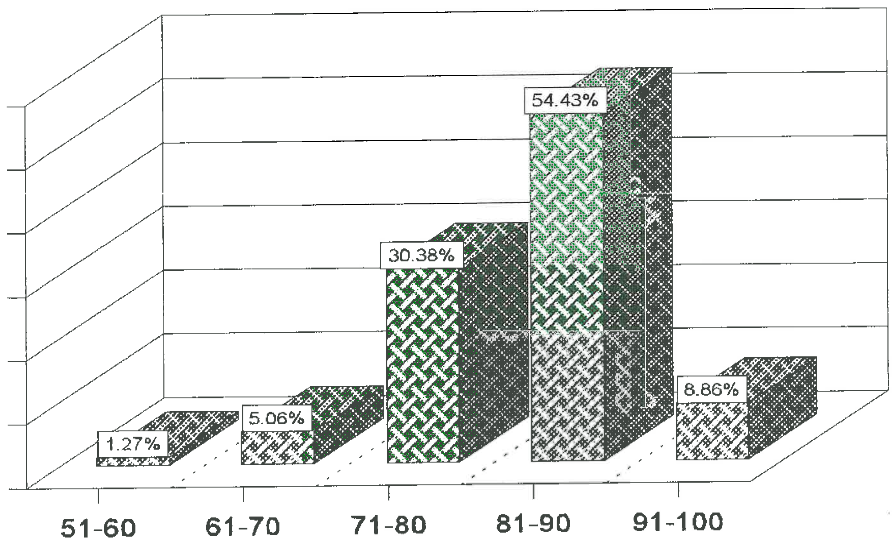


Figure 12.5: Frequency distribution of results obtained on the Numerical Ability index of the VPIB.

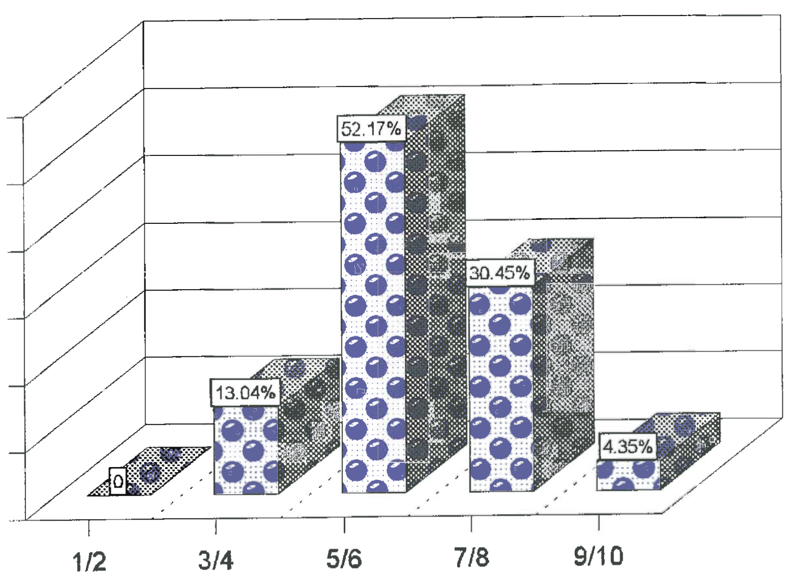


Figure 12.6: Frequency distribution of results obtained on the Composition of Wholes index of the VPIB.

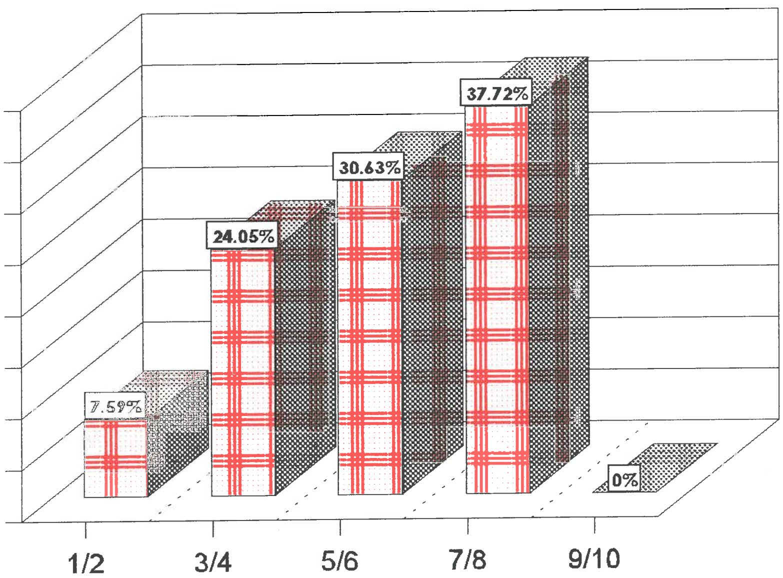


Figure 12.7: Frequency distribution of results obtained on the Spatial Reasoning index of the VPIB.

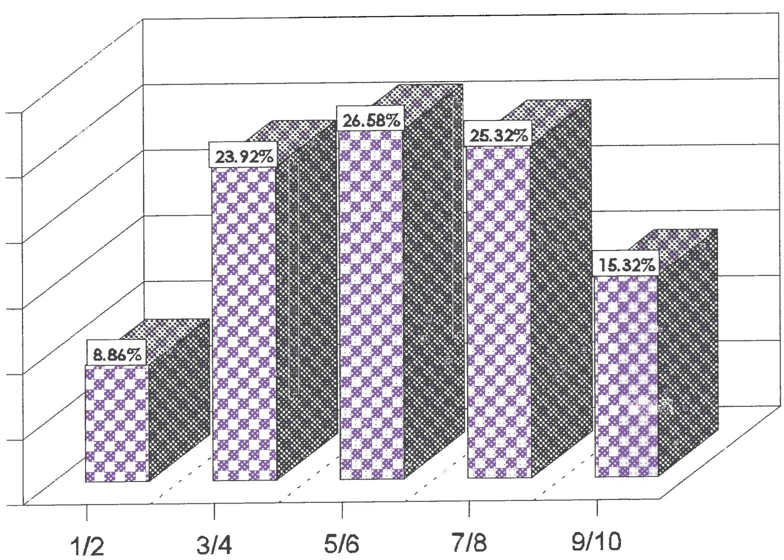
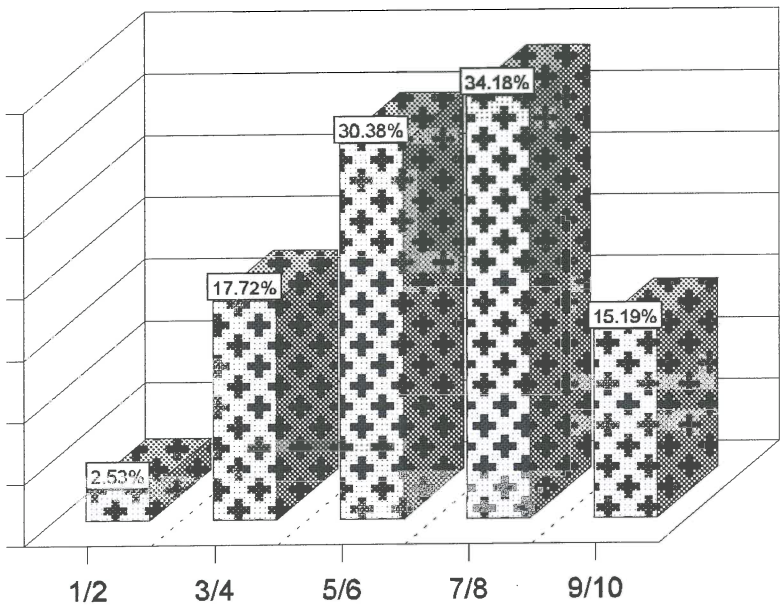


Figure 12.8: Frequency distribution of results obtained on the Perception index of the VPIB.



From Figure 12.5 it is clear that the majority of students admitted to the Engineering Technology courses obtained scores in the 5-6 range (the maximum score on this index equals 10) on the Numerical Ability index. Applicants who scored lower than 5 on this test were seldom admitted to the Engineering Technology courses, as it was felt that numerical ability should be important for Engineering Technology students. The smaller range is confirmed by the lower standard deviation calculated for this set of data, as depicted in Table 12.1.

Figure 12.6 shows that the bulk of the performance of students (68.35%) admitted to the engineering courses on the Composition of Wholes index fell into the 5-8 range. This index seemed to be fairly difficult, as none of the admitted applicants obtained a full score. Since applicants who scored low on this test were admitted in some instances, the standard deviation is higher than in the case of some of the other indices.

From Figures 12.7 and 12.8 it can be seen that applicants were admitted to the various Engineering Technology courses even though they performed poorly on the Spatial reasoning and Perception Indices. The majority of students obtained results in the 3 - 8 range.

12.3 Reliability of PIB indices

The following situation-specific reliability coefficients were obtained for the PIB-indexes used in the assessment of the academic potential of prospective Engineering Technology students at Technikon Pretoria:

When the possible answers to an item consisted of an item range, Cronbach's coefficient Alpha was computed. Where the possible answers were dichotomous, the Kuder-Richardson formula-20 was used.

Generally, a reliability coefficient of 0.75 for cognitive indices and 0.65 for emotional/social indices is considered acceptable.

Table 12.2: Reliability coefficients as computed for indices used in the assessment of the potential of prospective Engineering Technology students.

Index	Method	Reliability coefficient
Creativity	Cronbach	**0.75**
Reading Comprehension	KR20	**0.83**
Mental Alertness	KR20	**0.87**
Vocabulary	KR20	**0.76**
Numerical Ability	KR20	**0.68**
Composition of Wholes	KR20	**0.84**

Index	Method	Reliability coefficient
Spatial Reasoning	KR20	**0.79**
Perception	KR20	**0.79**

From Table 12.2 it can be seen that the reliability of the indices of the PIB fell into an acceptable range, except for that of the Numerical Ability index, which was lower than generally expected.

12.4 Predictive validity of variables

The final academic performance of first-semester Civil and Mechanical Engineering Technology students was used as the criterion for the calculation of the predictive validity coefficients. The predicate mark (year mark) was used as the final mark for students who did not write the final examination, or were not allowed (because of poor performance) to write the final examination. Firstly, the group was split into a Civil Engineering Technology and Mechanical Engineering Technology group, as some of the subjects included in the respective training programmes differ. Thereafter, the final average score for each student was calculated and used as criterion for the multiple regression analysis performed on the data. The results for each academic subject were then used as criterion for a multiple regression analysis performed. The results of the multiple regression analysis are given in Tables 12.3 to 12.15

Table 12.3: Result of multiple regression analysis performed on data, using a final average score for Civil Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
Multiple Regression Coefficient	0.319	0.432	0.537	0.594	0.604	0.614
R-Square	0.102	0.187	0.289	0.353	0.366	0.377
Standard Error of Estimate	8.986	8.566	8.028	7.670	7.610	7.558
F-Ratio	28.980	29.308	34.361	34.528	29.059	25.297
Degrees of Freedom	1,256	2,255	3,254	4,253	5,252	6,251
p-level	0.000	0.000	0.000	0.000	0.000	0.000
Constant	50.109	81.148	75.789	81.319	74.972	88.356
Mathematics	2.92	2.95	3.614	4.096	4.005	3.912
Creativity		-0.695	-0.773	-0.781	-0.735	-0.659
Composition of Wholes			1.146	1.423	1.326	1.483
Perception				-1.285	-1.303	-1.414
Numerical Ability					0.054	0.066
Mental Alertness						0.219

For the criterion variable ***Final Average Score*** for the Civil Engineering Technology course, R=0.614, and 37.7% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The Grade 12 symbol obtained for **Mathematics** (β=0.425) made the greatest contribution to the prediction of the final average score obtained by the students. However, adding the performance of the students on the **Composition of Wholes** index (β=0.410) of the Visual Potential Index Batteries to the prediction, the predictive validity of the assessment battery for the final average score obtained by Civil Engineering Technology students was significantly increased.

Students who scored low on Grade 12 Mathematics as well as on the Composition of Wholes index could thus have been expected to have a low final average score. The opposite would then also be true: a student who scored high on Grade 12 Mathematics as well as on the Composition of Wholes index could be expected to have a high final average score.

An inverted relation was found to exist between the criterion and the performance on the predictor variables **Creativity** and **Perception**. It thus seems as if those students who perceived themselves as less creative had a higher chance of achieving a high final average score than those who perceived themselves as innovative and not sticking to the tried ways. Furthermore, those students who were less able to perceive detail as well as wholes in their specific and logical context seemed to have performed better on the criterion variable Final Average Score.

The F Ratio in step 6 is statistically significant on the 0.1% level, thus indicating a significant relation between the actual and predicted scores on the criterion variable using the model incorporated in Table 12.3.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.49, and 24.5% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents $R = 0.46$, and 21.6% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

For the total group, the standard error of estimate = 7.558, which indicates that the prediction of performance is less accurate than expected. It is more accurate, however, than when only Mathematics and Physical Science symbols are used

(R=0.033; Standard error of estimate = 9.191). The standard error of estimate was possibly influenced by the fact that the Grade 12 symbols were used for the performance on Mathematics and Physical Science. A matric symbol only indicates a range of performance and the variance of performance within that range is disregarded.

The standard error of estimate varies if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 7.79 and for the Black group the standard error of estimate = 9.12. The percentage of Black students wrongly admitted as well as wrongly rejected, is 1.33% higher than the respective percentages for the White group. This could possibly be ascribed to the issue raised during the discussion of the literature review regarding the unreliability of school symbols obtained in former DET schools.

The question does arise whether it is worthwhile to include the final two steps in the prediction model. Two arguments influence this debate: (a) time and money are involved in the test administration process, which might be less if the battery were shortened; and (b) the state subsidy received by Technikon Pretoria is determined by the throughput rate of students. If the state subsidy should be increased to such an extent that the time and money spent, are justified by the 0.5% fewer applicants that will be wrongly accepted as well as wrongly rejected, , the full six steps should be included in the prediction model. If not, the final two steps should be omitted. In this instance, it will be acceptable to suffice with a four-step prediction model, where **$\hat{y}$= 81.319 + 4.096 (Mathematics) - 0.781 (Creativity) + 1.423 (Composition of Wholes) - 1.285 (Perception)**

Table 12.4: Result of multiple regression analysis performed on data, using a final Drawing score for Civil Engineering Technology, as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5
Multiple Regression Coefficient	0.452	0.589	0.634	0.689	0.693
R-Square	0.204	0.347	0.401	0.475	0.481
Standard Error of Estimate	11.003	10.018	9.633	9.061	9.058
F-Ratio	28.012	28.636	23.920	24.005	19.434
Degrees of Freedom	1,163	2,162	3,161	4,160	5,159
p-level	0.000	0.000	0.000	0.000	0.000
Constant	37.419	92.092	98.961	100.508	95.743
Composition of Wholes	2.194	2.206	2.451	2.698	2.624
Creativity		-1.175	-1.157	-1.302	-1.276
Perception			1.473	1.801	1.813
Mathematics				3.401	3.319
Numerical Ability					0.053

Table 12.4 shows that for the criterion variable ***Drawing*** for Civil Engineering Technology, R=0.693, and 48.1% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The most significant contribution to the prediction of performance on the criterion variable Drawing was the performance of the students on the **Composition of Wholes** index (β= 0.541)of the Visual Potential Index Batteries. It can therefore be deduced that the capacity for concrete reasoning as regards forms and figures and their logical place within a single structure plays an important role in the potential to master the more complicated field of Technical Drawing for Engineers.

By adding the students' performance on the predictor variable **Perception** (β=0.290), the predictive validity of the assessment battery is increased significantly. Those students who were able to perceive detail as well as wholes in their specific, logical context, seem to have a higher probability of being successful in Technical Drawing for Engineers.

The students' Grade 12 **Mathematics** performance further contributed significantly to the predictive validity of the assessment battery for the subject Technical Drawing for Engineers.

An inverted relation was found to exist between the criterion variable and the performance on the predictor variable **Creativity**. Again it seems as if those students who perceived themselves as being less creative had a better chance of achieving a high Drawing score than those who perceived themselves as innovative and not sticking to the tried ways.

It can therefore be expected that a student who scored high on the Composition of Wholes and Perception indices as well as on Grade 12 Mathematics would have a high Technical Drawing score. Furthermore a student who perceives himself or herself as less creative is more likely to obtain a high score on Technical Drawing.

The F Ratio in step 5 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.64, and 41% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents

R = 0.56, and 31.5% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 9.154. As in the case of the previous criterion variable, the standard error of estimate may be influenced by the fact that only symbols were available as an indication of the Mathematics performance in Grade 12, and the students' actual performances on this predictor variable could in fact, be anywhere within a particular range.

The standard error of estimate alters if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 8.13 and for the Black group the standard error of estimate = 13.80. The percentage of Black students wrongly admitted as well as wrongly rejected is 5.67% higher than that of the White group. Implemented in its current form, the prediction model might thus be biased for the criterion variable Drawing for Civil Engineering Technology and warrants further investigation.

The inclusion of the final step of the prediction model seems unnecessary, as it will not influence the accuracy of the model significantly. The final model is:
$\hat{y}$ = 100.508 + 2.698 (Composition of Wholes) -1.302 (Creativity) + 1.801 (Perception) + 3.401 (Mathematics)

From Table 12.5 it is clear that for the criterion variable ***Mathematics***, R = 0.807, and 65.1% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The Grade 12 symbol obtained for **Mathematics** (β= 0.652) made the greatest contribution to the prediction of the final Mathematics score obtained by the students. However, by adding the performance of the students on the

Composition of Wholes index (β= 0.572) of the Visual Potential Index Batteries to the prediction, the predictive validity of the assessment battery for the Mathematics score obtained by Civil Engineering Technology students was significantly increased. The addition of the performance of the students on the **Mental Alertness** index further enhanced the predictive validity of the assessment battery for the criterion variable Mathematics.

Table 12.5: Result of multiple regression analysis performed on data, using a final Mathematics score for Civil Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7
Multiple Regression Coefficient	0.475	0.550	0.648	0.744	0.761	0.784	0.807
R-Square	0.225	0.303	0.421	0.554	0.579	0.615	0.651
Standard Error of Estimate	10.958	10.424	9.528	8.386	8.169	7.831	7.477
F-Ratio	51.732	38.445	42.628	54.316	47.891	46.142	45.926
Degrees of Freedom	1,178	2,177	3,176	4,175	5,174	6,173	7,172
p-level	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Constant	41.618	52.033	44.834	113.175	108.639	122.422	133.154
Mathematics	5.517	6.107	7.046	7.183	6.790	6.951	7.674
Perception		-1.768	-2.509	-2.920	-2.844	-2.912	-2.594
Composition of Wholes			1.540	2.213	2.328	02.357	2.464
Mental Alertness				0.827	0.635	0.410	0.401
Reading Comprehension					-0.175	-0.242	-0.209
Creativity						-0.605	-0.811
Physical Science							-4.081

The relation between the criterion variable Mathematics and the matric symbol for **Physical Science** is reversed, as are the relations between the criterion variable and the performance on the predictor variables **Perception**, **Reading Comprehension** and **Creativity**.

It can be expected that a student who scored high on Grade 12 Mathematics, the Composition of Wholes and Mental Alertness indices will have a high Mathematic score.

It furthermore seems as if low performances on Perception, Creativity, Reading Comprehension and Grade 12 Physical Science can be associated with a high score in Mathematics for Civil Engineering Technology students.

The F Ratio in step 7 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.78, and 57.1% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents R = 0.67, and 47.4% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 7.477, which is again higher than would have been ideal, but still much lower than the standard error of estimate of 13.057 found for only the predictor variables Grade 12 symbol in Mathematics and Physical Science. Once again the issue of the use of symbols representing a range of possible performances already mentioned could explain the high standard error of estimate.

The standard error of estimate varies if the prediction model is applied to the main cultural groups separately. For the White group the standard error of estimate = 8.07 and for the Black group the standard error of estimate = 9.21. The percentage of Black students wrongly admitted as well as wrongly rejected is 1.14% higher than that of the White group. This could possibly again be ascribed to the unreliability of school symbols obtained in former DET schools.

As the addition of the final step increases by almost 4% the percentage of variance in the performance on the criterion variable explained, the seventh step is included in the prediction model.

The final model is: **$\hat{y}$= 133.154 + 7.675 (Mathematics) - 2.595 (Perception) + 2.464 (Composition of Wholes) + 0.401 (Mental Alertness) - 0.209 (Reading Comprehension) - 0.811 (Creativity) - 4.081 (Physical Science)**

Table 12.6: Result of multiple regression analysis performed on data, using a final Construction Materials score for Civil Engineering Technology as criterion.

	Step 1	Step 2	Step 3
Multiple Regression Coefficient	0.417	0.480	0.510
R-Square	0.174	0.230	0.260
Standard Error of Estimate	7.784	7.538	7.419
F-Ratio	32.035	22.614	17.536
Degrees of Freedom	1,152	2,151	3,150
p-level	0.000	0.000	0.000
Constant	44.969	68.267	52.978
Numerical Ability	0.224	0.211	0.186
Creativity		-0.488	-0.611
Mental Alertness			0.264

Table 12.6 indicates that for the criterion variable ***Construction Materials*** for Civil Engineering Technology, R = 0.510, and 26.0% of the variance in the criterion

variable can be explained by the variance in the predictor variables included in the prediction model.

The predictor variable **Numerical Ability** (β = 0.348) made the greatest contribution to the prediction of a final mark for the subject Construction Materials for Civil Engineering Technology. Those students who displayed the mainly verbal potential to calculate were thus those who had the greater possibility of being successful in the Construction Materials subject.

Performance on the **Mental Alertness** index (β = 0.187) made a slight contribution to the predictive validity of the assessment battery for the criterion variable Construction Materials.

An inverted relation between the criterion variable and the predictor variable **Creativity** was recorded.

A high score on the Numerical Ability and Mental Alertness indices is thus associated with a high score on the criterion variable Construction Materials, as is a low score on the Creativity index.

The F Ratio in step 3 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.50, and 25% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents, R = 0.54, and 31.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 7.419. As reported earlier in this chapter, a slightly lower than acceptable reliability coefficient was found for the Numerical Ability index. As reliability is known to influence validity, the error of estimate could partially be contributed to the limited reliability of the predictor variable.

The standard error of estimate varies if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 9.20 and for the Black group the standard error of estimate = 8.06. The percentage of White students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 1.14% higher than that of the Black group. It is noticeable that the school grades did not play a role in the prediction of the performance on this criterion variable, and it is the only model for Civil Engineering Technology where the standard error of estimate for the Black group is lower than for the White group. This might again confirm the deduction made through the literature study regarding the unreliability of school marks from former DET schools.

As the addition of the final step increases by 3% the percentage of variance in the performance on the criterion variable explained, the third step is included in the prediction model.

The final prediction model is: **$\hat{y}$ = 52.978 + 0.186 (Numerical Ability) -0.611 (Creativity) + 0.264 (Mental Alertness)**

Table 12.7 shows that for the criterion variable ***Applied Mechanics*** for Civil Engineering Technology, R = 0.536, and 26.0% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The Grade 12 symbols obtained for **Mathematics** (β= 0.303) and **Physical Science** (β= 0.285) made the greatest contribution to the prediction of the final Mechanics score obtained by the students.

The addition of the performances of the students on the **Composition of Wholes** (β= 0.247) and **Spatial Reasoning** (β= 0.270) indices of the Visual Potential Index Batteries to the prediction significantly increased the predictive validity of the assessment battery for the final Mechanics score obtained by Civil Engineering Technology students.

Inverted relations were found to exist between the performance on the criterion variable and the performances on the **Perception** and **Creativity** indices.

Students who scored low on Grade 12 Mathematics and Science as well as on the Composition of Whole and Spatial Reasoning indices could thus have been expected to have a low final Applied Mechanics score. High scores on the Perception and Creativity indices can also be associated with a high final Applied Mechanics score.

The F Ratio in step 6 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.54, and 29% of the variance in the criterion variable can be explained by the variance in the

Table 12.7: Result of multiple regression analysis performed on data, using a final Applied Mechanics score for Civil Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
Multiple Regression Coefficient	0.298	0.390	0.439	0.475	0.511	0.536
R-Square	0.089	0.152	0.193	0.225	0.261	0.287
Standard Error of Estimate	15.384	14.891	14.583	14.336	14.052	13.848
F-Ratio	14.161	12.933	11.375	10.320	9.953	9.406
Degrees of Freedom	1,145	2,144	3,143	4,142	5,141	6,140
p-level	0.000	0.000	0.000	0.000	0.000	0.000
Constant	43.415	57.234	52.751	46.273	78.324	82.425
Mathematics	4.450	4.504	4.461	5.472	4.628	4.860
Spatial Reasoning		0.918	0.853	0.879	0.713	0.723
Creativity			-2.117	-2.716	-2.867	-2.791
Physical Science				1.498	1.298	1.545
Composition of Wholes					3.547	4.946
Perception						-1.600

predictor variables included in the prediction model. For the Black respondents R = 0.40, and 16.3% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 13.848. Even though the standard error of estimate for this variable is higher than is acceptable, it is still lower than that of only the Grade 12 symbol in Mathematics and Physical Science, where the standard error of estimate = 15.613.

The standard error of estimate alters if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 11.03 and for the Black group the standard error of estimate = 13.7. The percentage of Black students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 2.67% higher than for the White group. As mainly the Grade 12 symbols for Mathematics and Physical Science contributed to the validity of the prediction model, the question raised during the literature study regarding the reliability of the results from former DET schools is raised again.

The inclusion of the final step increases the percentage of variance in the criterion variable with 2.5%, which is considered significant enough to make it worthwhile.

The final model is: **$\hat{y}$= 82.425 + 4.860 (Mathematics) + 0.723 (Spatial Reasoning) - 2.791 (Creativity) + 1.545 (Physical Science) + 4.946 (Composition of Wholes) - 1.600 (Perception)**

12.4.1 Summary of results found for Civil Engineering Technology

The results yielded by the multiple regression analysis performed on the data for Civil Engineering Technology indicate that specific predictor variables played a role in the prediction of the academic performance of students in more than one subject.

The Grade 12 Mathematics symbol consistently made the greatest contribution to the prediction of success in the Civil Engineering Technology programme. This finding corresponds with the general finding reported in Chapter 6, namely that school performance remains a good predictor of higher education academic performance. The Grade 12 Physical Science symbol does not perform as well as a valid predictor for success in the Civil Engineering Technology programme.

The addition of the performance on the Composition of Wholes index seems to enhance the predictive validity of the assessment battery for Civil Engineering Technology in almost all subjects.

Table 12.8: Result of multiple regression analysis performed on data, using a final average score for Mechanical Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7
Multiple Regression Coefficient	0.224	0.301	0.333	0.360	0.404	0.437	0.457
R-Square	0.050	0.090	0.111	0.130	0.163	0.191	0.209
Standard Error of Estimate	6.311	6.187	6.128	6.074	5.968	5.878	5.824
F-Ratio	14.480	13.567	11.308	10.111	10.523	10.592	10.102
Degrees of Freedom	1,274	2,273	3,272	4,271	5,270	6,269	7,268
p-level	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Constant	58.530	60.208	53.827	50.770	54.741	57.284	59.351
Creativity	-3.035	-5.423	-5.511	-5.691	-5.955	-6.125	-4.695
Composition of Wholes		1.291	1.243	0.942	0.985	1.052	0.803
Mathematics			0.138	0.134	0.117	0.117	0.105
Mental Alertness				0.837	1.547	2.268	2.676
Reading Comprehension					-1.407	-1.438	-1.300
Vocabulary						-1.943	-2.383
Spatial Reasoning							-2.163

For the criterion variable ***Final Average Score*** for the Mechanical Engineering Technology course, R=0.457, and 20.9% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The predictor variable **Mental Alertness** (β= 0.499) made the greatest contribution to the prediction of a final average score for Mechanical Engineering Technology. Those students who displayed the mainly verbal potential to classify objects correctly were thus those who had a greater possibility of obtaining a higher final average score. This test is also associated with the general mental ability or G-factor of intelligence (Schaap: 1997:70) and it could thus be said that general mental ability was the most significant predictor for the criterion variable final average score for Mechanical Engineering Technology in this study.

The addition of the performance of the students on the **Composition of Wholes** index (β= 0.166) of the Visual Potential Index Batteries to the prediction significantly increased the predictive validity of the assessment battery for the final average score obtained by the Mechanical Engineering Technology students.

A significant inverted relation was found to exist between the criterion and the performance on the predictor variable **Creativity**. It thus seems as if those students who perceived themselves as being less creative had a higher probability of achieving a high final average score than those who perceived themselves as innovative and not sticking to the tried ways.

Further significant inverted relations were also found between the criterion variable and the **Reading Comprehension** and **Vocabulary** indices of the PIB.

A student who achieved a high score on Mental Alertness and Composition of Wholes and low scores on Creativity, Reading Comprehension and Vocabulary had a good chance of obtaining a high final average score for Mechanical Engineering Technology.

The F Ratio in step 7 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.45, and 20.3% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents R = 0.51, and 30.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 5.82 which indicates that the predicted performance differs 5.82% on average from the observed performance. Using only these predictors may thus lead to a proportion of candidates admitted wrongly and the same proportion rejected wrongly.

The standard error of estimate varies if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 5.86 and for the Black group the standard error of estimate = 4.10. The percentage of White students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 1.76% higher than that of the Black group.

The final step in this prediction model seems redundant, as it does not make a significant contribution to the validity of the model.

The final prediction model is: **$\hat{y}$= 59.351 - 4.695 (Creativity) + 0.803 (Composition of Wholes) + 0.105 (Mathematics) + 2.676 (Mental Alertness) - 1.300 (Reading Comprehension) - 2.383 (Vocabulary)**

Table 12.9 indicates that for the criterion variable ***Electro-Technology*** for

Mechanical Engineering Technology, R = 0.559, and 31.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The predictor variable **Numerical Ability** made the greatest contribution to the prediction of a final mark for the subject Electro-technology for Mechanical Engineering Technology. Those students who displayed the mainly verbal potential to calculate were thus those who had the greater possibility of being successful in the subject Electro-Technology.

Table 12.9: Result of multiple regression analysis performed on data, using a final Electro-technology score for Mechanical Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7
Multiple Regression Coefficient	0.260	0.357	0.429	0.467	0.493	0.542	0.559
R-Square	0.069	0.127	0.184	0.218	0.243	0.293	0.312
Standard Error of Estimate	12.013	11.654	11.286	11.071	10.915	10.563	10.441
F-Ratio	20.323	19.890	20.488	18.886	17.301	18.613	17.370
Degrees of Freedom	1,274	2,273	3,272	4,271	5,270	6,269	7,268
p-level	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Constant	56.979	69.344	67.510	49.375	57.657	61.885	63.946
Numerical Ability	4.434	4.940	5.570	5.582	6.089	5.101	6.721
Reading Comprehension		-2.899	-5.025	-4.698	-4.955	-4.901	-5.504
Mental Alertness			3.070	3.127	3.888	2.920	2.704
Mathematics				0.283	0.315	0.273	0.308
Creativity					-4.469	-9.021	-12.380
Composition of Wholes						3.040	3.345
Perception							4.924

Performance on the **Composition of Whole** index (β= 0.360) made a significant contribution to the predictive validity of the assessment battery for the criterion variable Electro-Technology. The predictive validity of the assessment battery for this criterion variable was even further enhanced by adding the performance of the students on the **Mental Alertness** index (β= 0.262) of the PIB, as well as by the addition of the Grade 12 **Mathematics** symbol (β= 0.230).

There is a significant negative relation between the criterion variable and the performance on the predictor variables **Reading Comprehension** and **Creativity.**

A student who scored high on the Numerical Ability, Composition of Wholes as well as Mental Alertness indices and who obtained a high Grade 12 Mathematics symbol, and who scored low on the Reading Comprehension and Creativity indices can thus be expected to have a high Electro-Technology score.

The F Ratio in step 7 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.56, and 33.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents $R = 0.51$, and 29.6% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 10.441 which indicates that the predicted performance differs 10.44% on average from the observed performance. This implies a high percentage of applicants being rejected wrongly as well as the same percentage being admitted wrongly if only these predictors are used. It still is lower, however, than it would have been if only Grade 12 symbols for Mathematics and Physical Science had been used (standard error of estimate = 12.18). The standard error of estimate was possibly influenced by the fact that the Grade 12 symbols were used for the performance on Mathematics and Physical Science. A matric symbol only indicates a range of performance and the variance of performance within that range is disregarded.

The standard error of estimate varies if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 9.6 and for the Black group the standard error of estimate = 10.7. The percentage of Black students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 1.10% higher than that of the White group.

It is not worthwhile to include the final step of the prediction model, as it does not contribute significantly to the validity of the model.

The final model is: **$\hat{y}$= 63.946 + 6.721 (Numerical Ability) - 5.504 (Reading Comprehension) +2.704 (Mental Alertness) + 0.308 (Mathematics) -12.380 (Creativity) + 3.345 (Composition of Wholes)**

For the criterion variable ***Communication*** for Mechanical Engineering Technology R=0.600, and 36.0% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

Table 12.10: Result of multiple regression analysis performed on data, using a final Communication score for Mechanical Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
Multiple Regression Coefficient	0.366	0.451	0.541	0.561	0.583	0.600
R-Square	0.134	0.203	0.292	0.314	0.340	0.360
Standard Error of Estimate	7.533	7.238	6.834	6.739	6.623	6.533
F-Ratio	42.310	34.782	37.451	31.065	27.838	25.247
Degrees of Freedom	1,274	2,273	3,272	4,271	5,270	6,269
p-level	0.000	0.000	0.000	0.000	0.000	0.000
Constant	53.752	61.931	59.218	54.743	50.859	44.940
Composition of Wholes	2.210	2.587	1.809	2.167	2.324	1.597
Spatial Reasoning		-4.125	-5.911	-7.114	-7.725	-9.579
Mental Alertness			2.384	2.269	1.578	1.566
Reading Comprehension				1.889	2.471	3.076
Creativity					1.601	1.936
Numerical Ability						3.902

The predictor variable **Reading Comprehension** (β= 0.380) made the greatest contribution to the prediction of a final mark for the subject Communication for Mechanical Engineering Technology. Those students who displayed the competency to read and understand clearly what the reading matter conveyed, had a greater probability of being successful in the subject Communication than those who showed limited reading comprehension ability.

The addition of the performance of the students on the **Mental Alertness** index (β= 0.334) of the Potential Index Batteries to the prediction model significantly increased the predictive validity of the assessment battery for the final Communication score obtained by the Mechanical Engineering Technology students.

The performance of the students on the **Creativity** (β= 0.250) and **Composition of Wholes** (β= 0.264) indices enhanced even further the predictive validity of the assessment battery for the criterion variable Communication for Mechanical Engineering Technology.

A significant reversed correlation was found between the criterion variable and the predictor variable **Spatial Reasoning**.

A high score on the criterion variable Communication for Mechanical Engineering Technology is associated with high scores on the Reading Comprehension, Mental Alertness, Creativity and Composition of Wholes indices and low scores on the Spatial Reasoning index.

The F Ratio in step 6 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.58, and 33.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents R = 0.62, and 37.9% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 6.533, which means that 6.5% of the applicants would be respectively admitted or rejected wrongly if only these predictors were to be used.

The standard error of estimate varies if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 6.47 and for the Black group the standard error of estimate = 5.09. The percentage of White students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 1.38% higher than that of the Black group.

As the addition of the final step increases the percentage of variance in the performance on the criterion variable explained with almost 2%, the sixth step is included in the prediction model.

The model is: **$\hat{y}$= 44.940 + 1.598 (Composition of Wholes) - 9.579 (Spatial Reasoning) + 1.566 (Mental Alertness) + 3.076 (Reading Comprehension) + 1.936 (Creativity) + 3.902 (Numerical Ability)**

Table 12.11 shows that for the criterion variable ***Drawing*** for Mechanical Engineering Technology, R = 0.728, and 53.0% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The performance of students on the **Spatial Reasoning** index (β= 0.790) of the Visual Potential Index Batteries made the greatest contribution to the prediction of the final Drawing score obtained by the Mechanical Engineering Technology students. However, by adding the performance of the students on the **Numerical Ability** index (β= 0.581) of the Visual Potential Index Batteries to the prediction, the predictive validity of the assessment battery for the Drawing score obtained by Mechanical Engineering Technology students was significantly increased.

Table 12.11: Result of multiple regression analysis performed on data, using a final Drawing score for Mechanical Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7
Multiple Regression Coefficient	0.352	0.430	0.502	0.621	0.715	0.720	0.728
R-Square	0.124	0.185	0.252	0.386	0.511	0.518	0.530
Standard Error of Estimate	11.232	10.857	10.418	9.458	8.455	8.409	8.322
F-Ratio	38.855	30.930	30.555	42.558	56.437	48.191	43.127
Degrees of Freedom	1,274	2,273	3,272	4,271	5,270	6,269	7,268
p-level	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Constant	63.960	54.790	38.184	40.509	18.625	21.419	25.103
Composition of Wholes	3.158	3.677	3.158	2.049	4.755	4.598	4.987
Reading Comprehension		-2.907	-3.135	-4.686	-5.900	-6.226	-5.491
Numerical Ability			4.352	7.904	10.133	9.994	9.457
Spatial Reasoning				10.179	17.057	17.161	17.823
Creativity					-14.443	-14.534	-14.565
Vocabulary						-1.558	-2.988
Mental Alertness							1.844

The addition of the performance of the students on the **Composition of Wholes** index (β= 0.560) further enhanced the predictive validity of the assessment battery for the criterion variable Drawing.

Significant inverted relations were found between the performance on the criterion variable and the performances on the **Reading Comprehension** and **Creativity** indices.

A student who scored high on the Spatial Reasoning, Numerical Ability and Composition of Wholes indices, and who scored low on the Reading

Comprehension and Creativity indices, can thus be expected to have a high Drawing score.

The F Ratio in step 7 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.73, and 52.9% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents, R = 0.62, and 51.8% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 8.322 if the model is applied to the total group. The standard error of estimate alters marginally if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 8.018 and for the Black group the standard error of estimate = 8.410. The percentage of Black students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 0.39% higher than that of the White group.

The final step in this prediction model seems redundant, as it does not make a significant contribution to the validity of the model.

The final prediction model is: **$\hat{y}$= 21.419 + 4.598 (Composition of Wholes) - 6.226 (Reading Comprehension) + 9.994 (Numerical Ability) + 17.161 (Perception) - 14.534 (Creativity) - 1.558 (Vocabulary)**

Table 12.12: Result of multiple regression analysis performed on data, using a final Manufacturing Engineering score for Mechanical Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4
Multiple Regression Coefficient	0.646	0.780	0.805	0.808
R-Square	0.417	0.610	0.648	0.653
Standard Error of Estimate	6.291	5.156	4.910	4.885
F-Ratio	152.974	166.604	130.149	99.392
Degrees of Freedom	1,214	2,213	3,212	4,211
p-level	0.000	0.000	0.000	0.000
Constant	74.796	94.935	93.622	90.994
Spatial Reasoning	8.964	9.784	10.117	10.210
Creativity		-7.714	-11.110	-10.995
Numerical Ability			3.084	2.964
Reading Comprehension				0.781

For the criterion variable ***Manufacturing Engineering*** for Mechanical Engineering Technology R=0.808 and 65.3% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The predictor variable **Spatial Reasoning** (β= 0.740) made the greatest contribution to the prediction of a final mark for the subject Manufacturing Engineering for Mechanical Engineering Technology. Those students who displayed the competency to perceive, analyse and synthesise, as well as the ability to select and categorised were thus the students with the highest probability of being successful in the subject Manufacturing Engineering.

The addition of the performance of the students on the **Numerical Ability** index (β= 0.264) of the Visual Potential Index Batteries to the prediction model significantly increased the predictive validity of the assessment battery for the final Manufacturing Engineering score obtained by the mechanical engineering technology students.

A significant negative relation was found to exist between the criterion variable and the performance on the **Creativity** index.

A high score on the criterion variable Manufacturing Engineering for Mechanical Engineering Technology is associated with high scores on the Spatial Reasoning and Numerical Ability indices and low scores on the Creativity index.

The F Ratio in step 4 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.82, and 67.5% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents $R = 0.77$, and 59.6% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 4.89 if the prediction model is applied to the total group. The standard error of estimate alters noticeably if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 4.77 and for the Black group the standard error of estimate = 6.12. The percentage of Black students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 1.35% higher than that of the White group.

The final step in this prediction model seems to be redundant, as it does not make a significant contribution to the validity of the model.

The final model is: **$\hat{y}$= 90.994 + 10.210 (Spatial Reasoning) - 10.995 (Creativity) + 2.964 (Numerical Ability)**

Table 12.13 indicates that for the criterion variable ***Mechanics*** for Mechanical Engineering Technology, R = 0.451, and 20.3% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The performance of students on the **Mental Alertness** index (β= 0.320) of the Potential Index Batteries made the greatest contribution to the prediction of the final Mechanics score obtained by the Mechanical Engineering Technology students. However, by adding the Grade 12 **Mathematics** symbol (β= 0.313) to the prediction model, the predictive validity of the assessment battery for the Mechanical score obtained by Mechanical Engineering Technology students was significantly increased.

A significant inverted relation was found to exist between the criterion variable and the performance on the **Creativity** index.

A student who scored high on the Mental Alertness index as well as on Grade 12 Mathematics and who scored low on the Creativity index can thus be expected to have a high Mechanics score.

Table 12.13: Result of multiple regression analysis performed on data, using a final Mechanics score for Mechanical Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5
Multiple Regression Coefficient	0.292	0.412	0.421	0.436	0.451
R-Square	0.085	0.170	0.178	0.190	0.203
Standard Error of Estimate	15.771	15.053	15.010	14.922	14.827
F-Ratio	25.610	27.941	19.585	15.915	13.796
Degrees of Freedom	1,274	2,273	3,272	4,271	5,270
p-level	0.000	0.000	0.000	0.000	0.000
Constant	2.073	22.547	27.356	25.675	31.979
Mathematics	0.588	0.642	0.626	0.656	0.630
Mental Alertness		10.067	9.454	10.668	10.985
Vocabulary			-2.074	-4.060	-4.135
Creativity				-2.056	-3.245
Reading Comprehension					-2.265

The F Ratio in step 5 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.49, and 23.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents, R = 0.43, and 19.3% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 14.827 if the prediction model is applied to the total group, which indicates that the predicted performance differs 14.83% on average from the observed performance. This implies a high percentage of applicants being rejected wrongly as well as the same percentage being admitted wrongly if only these predictors are used. It still is lower, however, than it would have been if only Grade 12 symbols for Mathematics and Physical Science were used (standard error of estimate = 15.771). Again, the issue of the use of symbols representing a range of possible performances already mentioned could explain the high error of estimate.

The standard error of estimate alters noticeably if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 13.58 and for the Black group the standard error of estimate = 16.92. The percentage of Black students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 3.34% higher than for the White group. As the Grade 12 symbol for Mathematics contributed mainly to the validity of the prediction model, the question raised during the literature study regarding the reliability of the results from former DET schools arises again.

It is not worthwhile to include the final step of the prediction model, as it does not contribute significantly to the validity of the model.

The final prediction model is: **$\hat{y}$= 25.675 + 0.656 (Mathematics) + 10.668 (Mental Alertness) - 4.060 (Vocabulary) - 2.056 (Creativity)**

Table 12.14: Result of multiple regression analysis performed on data, using a final Computer Skills score for Mechanical Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
Multiple Regression Coefficient	0.361	0.537	0.552	0.578	0.621	0.672
R-Square	0.131	0.288	0.305	0.334	0.386	0.451
Standard Error of Estimate	9.451	8.568	8.482	8.317	7.999	7.580
F-Ratio	41.146	55.253	39.751	33.991	33.981	36.827
Degrees of Freedom	1,274	2,273	3,272	4,271	5,270	6,269
p-level	0.000	0.000	0.000	0.000	0.000	0.000
Constant	52.397	60.786	59.399	79.629	86.692	83.264
Mental Alertness	3.033	5.424	5.367	4.622	4.812	4.250
Reading Comprehension		-4.746	-5.080	-5.047	-5.203	-5.536
Composition of Wholes			0.286	0.322	0.314	0.286
Creativity				-1.443	-2.476	-3.083
Spatial Reasoning					4.602	7.980
Numerical Ability						-4.974

For the criterion variable ***Computer Skills*** for Mechanical Engineering Technology R=0.672 and 45.1% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The greatest contribution to the prediction of a final mark for the subject Computer Skills for Mechanical Engineering Technology was made by the predictor variable **Mental Alertness** (β= 0.579). The general mental ability associated with performance on this index, thus seems to have influenced the performance of students in the subject Computer Skills.

The addition of the performance of students on the **Spatial Reasoning** index (β= 0.499) further enhanced the predictive validity of the assessment battery for the criterion variable Computer Skills. Those students who displayed the competency to perceive, analyse and synthesise, as well as the ability to select and categorise were thus the students with the highest probability of being successful in the subject Computer Skills. The addition of the performance on the **Composition of Wholes** index (β= 0.345) further significantly contributed to the prediction model.

Significant negative relations were found to exist between the criterion variable and performances on the predictor variables **Creativity, Reading Comprehension** and **Numerical Ability**.

A high score on the criterion variable Computer Skills for Mechanical Engineering Technology is associated with high scores on the Mental Alertness, Spatial Reasoning and Composition of Wholes indices and low scores on the Creativity, Reading Comprehension and Numerical Ability indices.

The F Ratio in step 6 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.66, and 44.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents, R = 0.69, and 47.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 7.58 if the prediction model is applied to the whole group, which, even though high, is still lower than it would have been if only Grade 12 performance in Mathematics and Science were used.

The standard error of estimate alters marginally if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 7.65 and for the Black group the standard error of estimate = 7.39. The percentage of White students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 0.26% higher than that of the Black group.

As the addition of the final step increases with almost 6.5% the percentage of variance in the performance on the criterion variable explained, the sixth step is included in the prediction model.

The final model is: **$\hat{y}$= 83.264 + 4.250 (Mental Alertness) - 5.536 (Reading Comprehension) + 0.286 (Composition of Wholes) - 3.083 (Creativity) + 7.980 (Spatial Reasoning) - 4.974 (Numerical Ability)**

Table 12.15 shows that for the criterion variable ***Mathematics*** for Mechanical Engineering Technology, R = 0.680, and 46.3% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model.

The Grade 12 symbol obtained for **Mathematics** (β= 0.558) made the greatest contribution to the prediction of the final Mathematics score obtained by the students. However, by adding the performance of the students on the **Numerical Ability** index (β= 0.200) of the Visual Potential Index Batteries to the prediction, the predictive validity of the assessment battery for the Mathematics score obtained by Mechanical Engineering Technology students was significantly increased.

A significant inverted relation was found to exist between the performance on the criterion variable and the performances on the **Reading Comprehension** index.

Table 12.15: Result of multiple regression analysis performed on data, using a final Mathematics score for Mechanical Engineering Technology as criterion.

	Step 1	Step 2	Step 3	Step 4	Step 5
Multiple Regression Coefficient	0.587	0.636	0.667	0.671	0.680
R-Square	0.345	0.404	0.444	0.450	0.463
Standard Error of Estimate	10.122	9.674	9.357	9.329	9.253
F-Ratio	144.326	92.508	72.523	55.377	38.592
Degrees of Freedom	1,274	2,273	3,272	4,271	5,270
p-level	0.000	0.000	0.000	0.000	0.000
Constant	-11.563	3.025	-9.281	-7.545	-6.964
Mathematics	0.897	0.834	0.834	0.856	0.826
Reading Comprehension		-2.950	-2.650	-2.295	-2.893
Numerical Ability			3.435	4.058	4.029
Spatial Reasoning				-2.024	-2.606
Mental Alertness					1.016

A student who scored high on Grade 12 Mathematics and on the Numerical Ability index and who scored low on the Reading Comprehension index can thus be expected to have a high Mathematics score.

The F Ratio in step 5 is statistically significant on the 0.1% level, thus indicating a significant relation between the performance on the predictor variables and the criterion variable.

If the prediction model is applied to the major cultural groups separately, the multiple regression coefficient for the White respondents equals 0.71, and 48.9% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For the Black respondents, $R = 0.63$, and 41.2% of the variance in the criterion variable can be explained by the variance in the predictor variables included in the prediction model. For both groups, the F value remains statistically significant on the 0.1% level.

The standard error of estimate = 9.25 for the total group, which indicates that the predicted performance differs 9.25% on average from the observed performance. This implies a high percentage of applicants being rejected wrongly as well as the same percentage being admitted wrongly if only these predictors were to be used. It still is lower, however, than it would have been if only Grade 12 symbols for Mathematics and Physical Science were used (standard error of estimate = 10.12). The high error of estimate could again be explained by the issue of the use of symbols representing a range of possible performances already mentioned.

The standard error of estimate alters marginally if the prediction model is applied to the main cultural groups separately. For the White group, the standard error of estimate = 8.25 and for the Black group the standard error of estimate =10.78. The percentage of Black students that would be wrongly admitted as well as wrongly rejected if this model were to be implemented unchanged, is 2.53% higher than that of the White group. As the Grade 12 symbol for Mathematics contributed mainly to the validity of the prediction model, the question raised during the literature study regarding the reliability of the results from former DET schools, arises once again.

The final two steps of the model do not contribute significantly to the validity of the prediction model and can thus be omitted.

The final model is: **$\hat{y}$= −9.281 + 0.834 (Mathematics) - 2.650 (Reading Comprehension) + 3.435 (Spatial Reasoning)**

12.4.2 Summary of results found for Mechanical Engineering Technology

From the above it is clear that specific predictor variables played a significant role in the prediction of performance for more than one subject in the Mechanical Engineering Technology course.

The Grade 12 Mathematics symbol consistently made a contribution to the prediction of success in the Mechanical Engineering Technology programme. This finding corresponds with the general finding reported in Chapter 6, namely, that school performance remains a good predictor of higher education academic performance. The Grade 12 Physical Science symbol does not perform as well as a valid predictor for success in the Mechanical Engineering Technology programme.

The addition of the performance on the Numerical Ability index seems to enhance the predictive validity of the assessment battery for Mechanical Engineering Technology in at least four subjects. In view of the fact that the reliability of this index was shown earlier to be lower than expected, its predictive validity would possibly be higher if the reliability could be improved.

The performance of the students on the Mental Alertness index further added to the predictive validity of the assessment battery, as it showed significant correlation with the final average score as well as with various individual subjects in the Mechanical Engineering Technology programme.

The Spatial Reasoning index correlated significantly with some subjects, but reversed relations with other were recorded.

12.5 Conclusion

The results of the statistical analysis performed on the data were presented in this chapter. It seems as if the proposed assessment battery has a higher predictive validity than has the use of the Grade 12 Mathematics and Physical Science symbols only for selection purposes. The conclusions drawn from these results are presented in Chapter 13, together with specific recommendations regarding the implementation of the potential assessment battery.

Chapter 13

Conclusions and recommendations

13.1 Introduction

Since the results of the study were presented and discussed in the previous chapter, this chapter contains the conclusions drawn from the literature study and research findings. Furthermore recommendations will be made towards future research and the possible development and refining of the suggested assessment battery.

13.2 Conclusions

This study addressed the need for a valid, accurate and efficient tool for the selection of Engineering Technology students. From the research findings and literature study the following deductions regarding the suggested potential assessment battery can be made:

13.2.1 The results of the study revealed that the accurate selection of Engineering Technology students is possible to a satisfactory extent.

13.2.2 The expansion of the traditional selection procedure to include the potential assessment phase proved valuable, as the validity of all prediction models improved with the addition of the indices from the Potential Index Batteries. The prediction models were found to be unbiased against students from the previously disadvantaged school systems and can thus be said to be culture fair.

The incorporation of the suggested potential assessment battery in the selection process will thus contribute to increased and broadened participation in higher education as prescribed by the White Paper on Higher Education (DOE, 1997:8) and confirmed by the National Plan for Higher Education (Ministry of Education, 2001:15), as it will give students who were previously excluded on the grounds of school performance an opportunity to be included in higher education . The development of alternative selection methods that make higher education more accessible for different race groups is therefore in accordance with the guidelines set for the transformation of higher education by the Department of Education (DOE, 1997:8). Such a potential assessment system will furthermore support the vision of the Ministry of Education to develop a system of higher education that will promote equity of access, while eradicating most, if not all, forms of unfair discrimination, as discussed in Chapter 2.

Empirical

13.2.3 The results of the statistical analysis performed on the research data proved that not all the indices included in the potential assessment process made a significant contribution to the prediction of academic performance in the Engineering Technology courses. However, some of the indices of the Potential Index Batteries did make consistent contributions to the prediction of performance in the Engineering Technology courses. Table 13.1 contains a matrix that indicates which variable predicted performance on which criterion variable.

13.2.5 It is also clear from Table 13.1 that different indices from the Potential Index Batteries played a significant role in the prediction of performance in the Civil Engineering and Mechanical Engineering Technology courses, respectively. This finding confirms the concept of situation-specific selection, as proposed by the developers of the Potential Index Batteries and discussed in Chapter 7.

13.2.6 The Potential Index Batteries were found to be both reliable and valid in general. It was furthermore proved that they were not biased against any population group included in this study.

13.2.7 A disturbingly high standard error of estimate was found throughout the study. This could possibly be ascribed to the unreliability of the criterion scores used in the study, which might be explained by the fact that the academic results of students were obtained over a three-year period.

13.2.8 This study does not claim to say the final word for the selection of Engineering Technology students. It has however strengthened the validity and fairness of the selection procedures and increased the possibility for selection transparency. The research indicated the areas in which the selection mechanism was valid, but also the aspects of it that were not justifiable.

13.3 Recommendations

The following recommendations are made from the conclusions that have been drawn:

13.3.1 Even though the Grade 12 symbol for Mathematics of the students from the former DET schools did not have as high a predictive validity as had been expected, the concept of prior knowledge of Mathematics, as predictor of academic success in engineering courses, and as represented by the symbol obtained in the final high school examination, should not be negated. The fact that the current matriculation system will be replaced by the Further Education Training Certificate (FECT) also has an impact. The exact format of the final results of this system is currently unknown and the predictive validity of the new end-of-school results will have to be investigated before they can be used for selection purposes. This situation emphasises the need for higher education institutions to develop and validate their own instruments and policies. It is recommended that the initiative to develop an in-house Mathematics and Science admission test, as Technikon Pretoria is currently doing, should be investigated further and that the reliability and validity of such an admission test be established empirically. This admission test should, in conjunction with the potential assessment battery suggested by this study, form the core of the selection procedure for Civil and Mechanical Engineering Technology students at Technikon Pretoria.

13.3.2 It is recommended that the potential assessment battery for Civil Engineering Technology should consist at least of Creativity, Mental Alertness, Numerical Ability, Composition of Wholes and Perception. The battery for Mechanical Engineering Technology should at least include Creativity, Reading Comprehension, Mental Alertness, Numerical Ability, Composition of Wholes and Spatial Reasoning.

13.3.3 As an alternative, it is proposed that all prospective Engineering Technology students be evaluated with the same battery and that different weights be allocated to different indices, as determined by the prediction models. It would then be possible to implement an advisory selection system, where the most suitable course for a candidate, according to his or her potential profile, could be recommended.

13.3.4 It is furthermore recommended that this study be repeated for the Electrical and Industrial Engineering Technology courses in order to establish prediction models for those courses, to ensure that the

	Creativity	Reading Comprehension	Mental Alertness	Vocabulary	Numerical Ability	Composition of Wholes	Spatial Reasoning	Perception	Mathematics	Physical Science
Final Average (C)	X		X		X	X		X	X	
Drawing (C)	X					X		X	X	
Mathematics (C)	X	X	X			X			X	X
Construction Materials (C)	X				X					
Applied Mechanics (C)	X					X	X	X	X	
Final Average (M)		X	X	X		X			X	
Electro-Technology (M)	X	X	X		X	X	X	X	X	
Communication Skills (M)	X	X	X			X	X			
Drawing (M)	X	X		X	X	X	X			
Manufacturing Engineering (M)	X				X					
Mechanics (M)	X		X	X					X	
Computer Skills (M)	X	X	X		X	X	X			
Mathematics (M)		X			X				X	

C = Civil Engineering Technology
M = Mechanical Engineering Technology

Table 13.1: Matrix of variables that played a role in the prediction of performance on the criterion variables

advisory selection process functions optimally.

13.3.5 The possibility of expanding the potential assessment battery to include other predictors that were not included in this study should be investigated.

13.3.6 The reliability and consistency of the criterion scores used in this study are questioned and the way in which marks are allocated in the various academic subjects should be investigated.

13.3.7 If the Potential Assessment Battery is to be used in a selection context where the characteristics of the applicant pool differ from the sample used here, or where the programme for which students are selected differs significantly from the Engineering Technology courses used as criterion variables in this study, the effectiveness and fairness of the potential assessment test will have to be re-evaluated for the new context.

13.3.8 Since predictions involving short time spans (e.g. weeks or months) usually are more accurate than those where there is a long time span between the assessment of the predictor and criterion variables (Mc Millan & Schumacher, 1993:275) it is recommended that a study is undertaken to monitor the eventual success of students admitted by means of the selection procedure suggested here.

13.4 Limitations of this study

The execution of this study posed specific limitations that are discussed below.

13.4.1 No aptitude or potential test will ever correlate perfectly with academic performance, because motivation and effort always play a role as well (Huysamen, 1997:68). The validity of the prediction model was most likely negatively affected by the fact that these factors were not taken into consideration or included in the contingent of possible predictors,

13.4.2 As only students who had already been admitted to the various Engineering Technology courses were included in the sample, the study was performed on a pre-selected group. The range of scores on the predictor variables was confined to a representation of only a part of the total distribution of that variable, and this restriction in range could have led to a lowering of the correlation. McMillan and Shumacher (1993:277) state that restriction in range partially explains the usually modest relationship between college admission tests and academic achievement - the range is restricted to those students who achieved high scores in the admission tests.

13.4.3 It could be argued that academic performance is not the only relevant measurement when it comes to the selection of students. Issues such as performance on the actual job could also be regarded as relevant. The counter argument could be that a student who will not be successful in the job market should not pass the course.

As this study was initiated to address the specific problem of admitting students to higher education, academic success was deemed criterion enough. Future studies could focus on the predictive validity for performance in the labour market of the suggested selection battery used in this study .

13.4.4 Since a nonprobability convenience sample was selected for this study, the results could not have been generalised from sample to population, as the sampling error could not be determined. The results of this study are thus valid only for the selection of Civil and

Mechanical Engineering Technology students at Technikon Pretoria.

List of References

African National Congress. 1994. The Reconstruction and Development Programme: A policy framework. Johannesburg: Umyano Publications.

American Educational Research Association, American Psychological Association & National Council on Measurement in Education.1985. Standards for educational and psychological testing. Washington, DC: American Psychological Association.

Anastasi, A. & Urbina, S. 1997. Psychological testing. 7th edition. New York: Macmillan Publishing Company.

Armstrong, M. 1996. A Handbook of Personnel Management Practice. 6th edition. London: Kegan Page Limited.

Arvey, R.D. & Sackett, P.R. 1993. Fairness in selection: Current Developments and Perspectives. In: Schmitt, N.; Borman, W.C. and Associates. Personnel Selection in Organizations. San Francisco: Jossey-Bass Publishers.

Atlink, W.M.M. 1987. The evaluation of selection tests for educational upgrading programs in Botswana and Swaziland. International Journal of Educational Development, 7(1). 1-12.

Bingham, W.V. 1937. Aptitudes and aptitude testing. New York: Harper & Row.

Blake, R.H. 1983. Personnel Selection. In: Barling, B (ed.) Behaviour in Organisations: South African Perspectives. Johannesburg: McGraw-Hill.

Bless, C. &Kathuria, R. 1993. Fundamentals of Social Statistics: An African Perspective. Cape Town: Wetton; Johannesburg: Juta and Co. Ltd.

Bodibe, R.C. 1997. Access, throughput and output maps that no longer match the terrain. Paper presented at the 17th Annual Conference of the Society for Student Counselling in Southern Africa, Technikon Port Elizabeth.

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 for Higher Education, 6(1). 59-65.

Brown, F.G. 1983. Principles of Educational and Psychological testing. 3rd edition. USA: Holt, Rinehart and Winston.

Brusselmans-Dehairs, C. & Henry, G.F. 1994. Introduction to special issue on gender and mathematics. International Journal of Educational Research, 21(4), 353-371.

Cascio, W.F. 1991. Applied Psychology in Personnel Management. 4th edition. Englewood Cliffs: Prentice-Hall.

Cattell, R.B. & Eber, W.H. 1961. Handbook for the Sixteen Personality Factor Questionnaire. Campaign: Institute for Personality and Ability Testing.

Cole, N.S. 1973. Bias in selection. Journal of Educational Management, 10. 237-255.

Committee of University Principals. 1987. Macro-aspects of the university within the context of tertiary education in the Republic of South Africa. Report of the Main Committee of the CUP: Pretoria.

Constitution of the Republic of South Africa. 1996. Act 108.

Cronbach, L.J. 1990. Essentials of psychological testing. 3rd edition. New York: Harper and Row.

Dane, F.C. 1990. Research methods. California: Brooks/Cole.

Darlington, R.B. 1971. Another look at "culture fairness." Journal of Educational Measurement, 8. 71-82.

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 for Student Counselling in Southern Africa, Technikon Port Elizabeth.

De la Rey, R.P. 1976. Sielkundige hoedanighede van en prestasiemaatstawwe by MBA-studente. Unpublished PhD Thesis: Pretoria: University of Pretoria.

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.

Einhorn, H.J. & Bass, A.R. 1971. Methodological considerations relevant to discrimination in employment testing. Psychological Bulletin, 75. 261-269.

Erasmus, P.F. 1997. Questions of Standardisation, Validity and Reliability of Assessment Tools. In: P.F. Erasmus(Ed). The Unfair Labour Practice the Worker and Assessment. Krugersdorp: RNATA.

Erasmus, P.F. & Minnaar, G.G. 1995. Manual for Erasmus/Minnaar Potential Index (Intermediate Scale). Krugersdorp: Potential Index Associates.

Fourie, C.M. 1988. 'n Keuringsmetode vir eerstejaar-ingenieurswesestudente. South African Journal for Higher Education, 2 (2). 73-77.

Geldenhuys, E.G. 1996. 'n Ondersoek na die differensiële funksionering van die "Raven's standard progressive matrices" ten opsigte van blanke en swart Suid-Afrikaners. Unpublished Thesis: M Com (Human Resource Management): University of Pretoria.

Gottfredson, L.S. 1986. Validity versus Utility of Mental Tests: Example of the SAT. Journal of Vocational Behavior, 29. 363-378.

Gregory, R.J. 1996. Psychological testing: history, principles and applications. Boston: Allyn and Bacon.

Greuter, M.A. 1989. Performance Modelling for Personnel Selection. In: Herriot, P. (ed). Assessment and Selection in Organizations: Methods and Practise for Recruitments and Appraisal. Chichester, England; New York: Wiley.

Guy, R.F.; Edgley, C.E.; Arafat, I. &Allen, D.E. 1987. Social Research Methods: Puzzles and Solutions. Boston; London; Sydney; Toronto: Allyn and Bacon, Inc.

Hair, J.F. ; Anderson; Totham &Black.1998. Multivariate Data Analysis with Readings. Upper Saddle River, N.J.: Prentice-Hall.

Hessels, M.G.P. 1995. Leertest voor Etnische Minderen: Teoretische en

Empirische Verantwoording. RISBI/Swets en Zeitlinger, Thuispagina Erasmus Universiteit, Rotterdam.

Howell, D.C. 1992. Statistical methods for psychology. 3rd edition. California: Duxbury Press.

Human Sciences Research Council. 1995. Ways of seeing the National Qualifications Framework. Pretoria: Human Sciences Research Council.

Huysamen, G.K. 1996. Fair and unbiased admission procedures for South African institutions of higher education. South African Journal for Higher Education, 10 (2). 199-207.

Huysamen, G.K. 1997. Potential ramifications of admission testing at South African institutions of higher education. South African Journal for Higher Education, 11 (1). 65-71.

Jacobs, G.J. 1987. 'n Voorgestelde metode by die keuring van universiteitstudente. Bulletin vir Dosente, 19 (2). 25-40.

Kline, P. 1993. The Handbook of Psychological Testing. London and New York: Routledge.

Kotze, H.N. 1994. Keuringsmodelle vir Universiteitstudierigtings: 'n Psigometriese ondersoek. Unpublished PhD Thesis: Potchefstroom University for Christian Higher Education.

Kotze,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 for Student Counselling in Southern Africa, PREMOS, Pretoria.

Latham, G.P. &Wexley, K.N. 1981. Increasing Productivity Through Performance Appraisal. Massachusetts, California, London: Addison-Wesley Publishing Company Inc.

Leap,T.L. &Crino, M.D. 1993. Personnel/Human Resource Management. 2nd edition. United States of America. Macmillan Publishing Company.

Louw, A.D. 1992. Selection of prospective students for tertiary education: A creative response to the turbulence and diversity of the nineties. Paper presented at the 12th Annual Conference of the Society for Student Counselling in Southern Africa, Cape Town.

McMillan, J.H. & Schumacher, S. 1993. Research in Education: A Conceptual Introduction. 3rd edition. New York: Harper Collins College Publishers.

Mehrens, W.A. &Lehmann, I.J. 1991. Measurement and evaluation in Education and Psychology. 4th edition. USA: Harcourt Brace Jovanovich College Publishers.

Ministry of Education 2001. National Plan for Higher Education.

Moody,C.D. 1994. Strategies for improving the representation and participation of Black faculty, students and staff in tertiary education. Paper presented at Conference on Multiculturism, Vaal Triangle Technikon.

Moran, V.C. 1987. Justifying admission criteria. Unpublished preliminary report. University of Cape Town.

Murphy, K.R. & Davidshoffer, C.O. 1994. Psychological Testing: Principles and

Applications. 3rd edition. Colorado: Prentice-Hall International, Inc.

National Union of Metal Workers in South Africa. 1992. Code of practice to end unfair discrimination in South Africa. Unpublished document. NUMSA.

Nzimande, B. 1995. To test or not to test? Paper presented at the Congress on Psychometrics, CSIR, Pretoria.

Pienaar, J.L. 1991. Die voorspelling van akademiese sukses met behulp van die AAT. Unpublished Thesis: VISTA University. Pretoria

Pieters, M.A. 1996. Human Resource Practitioners. Pretoria: Kagiso Publishers.

Popham, W.J. & Sirotnik, K.A.1992. Understanding Statistics in Education. Illinois: F.E. Peacock Publishers, Inc.

Pottas, C.D.; Erwee, R; Boshoff, A.B. & Lessing, B. 1988. Manual for the achievement motivation test. Pretoria: Unit for entrepreneurship, University of Pretoria.

Roe, R.A. 1989. Designing Selection Procedures. In: Herriot, P. (ed). Assessment and Selection in Organizations: Methods and Practice for Recruitments and Appraisal. Chichester, England; New York: Wiley.

Roe, R.A.& Greuter, M.A. 1991. Developments in personnel Selection Methodology. In: Hambleton, R.K. and Zaal, J.N. Advances in Educational and Psychological Testing: Theory and Applications. Boston/ London/Dortrecht: Kluwer Academic Publishers.

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 to the first-year Accounting programme - preliminary evidence from the university of Durban-Westville. South African Journal of Education, 16 (2). 117-122.

Sax, G. 1980. Principles of Educational and Psychological measurement and evaluation. 2nd edition. Belmont: Wadsworth Publishing Company.

Schaap, P. 1995. Psigometrika: Aanvullende en Kernaantekeninge. Unpublished handout. University of Pretoria.

Schaap, P. 1997. Guidelines for Evaluating and Validating Assessment Tools. In: Erasmus, P.F. (Ed) The Unfair Labour Practice the Worker and Assessment. Krugersdorp: RNATA.

Schoombee, C.J. 1992. Die geldigheid van maatstawwe vir die keuring van beurshouers binne 'n nutsmaatskappy. Unpublished MCom Thesis: University of Pretoria.

Shirley, D.W. 1994. Psychometric testing and organisational culture: Implications for affirmative Action. Paper presented at the Congress on Psychometrics for Psychologists and Personnel practitioners: Society for Industrial Psychology of South Africa.

Skuy, M.; Zolezzi, S.; Mentis, M.;Frijdhorn, P. & Cockcroft, K. 1996. Selection of advantaged and disadvantaged South African students for university admission. South African Journal for Higher Education,10 (1). 110-117.

Smit, G.J. 1983. Navorsingsmetodes in die Gedragswetenskappe. Pretoria: Gutenberg Boekdrukkers.

Smit, G.J. 1991. Psigometrika: Aspekte van toetsgebruik. 2nd edition. Pretoria: HAUM.

Smit, G.J. 1996. Psychometrics: Aspects of measurement. Pretoria: Kagiso Publishers.

Smit, P.C. 1983. Skoolprestasie as voorspeller van akademiese sukses van leerlingingenieurstegnici. Pretoria: Human Sciences Research Council.

Smith, E.H. 1992. An outline for implementing or improving selection procedures for prospective technikon students, with specific references to engineering courses. Paper presented at the 12th Annual Conference of the Society for Student Counselling in Southern Africa, Cape Town.

Smith, M. & Robertson, I.T. 1986. The theory and practice of systematic staff selection. London: Macmillan Press Ltd.

Snyman, J.J. 1987. Ondersoek na moontlike toelatingsvereistes tot Regstudie. Unpublished research report: Rand Afrikaans University.

Society for Industrial and Organizational Psychology. 1987. Principles for the validation and use of personnel selection procedures. College Park, MD: Author.

Stoker, D.J.; Engelbrecht, C.S.; Crowther, N.A.S.; Du Toit, S.H.C. & Herbst, A. 1984. Ondersoek na differensiële toelatingsvereistes tot tersiêre onderwysinrigtings. Pretoria: Human Sciences Research Council.

Stumph, R. 1997. The White Paper on Higher Education: Implications for counselling and guidance of students particularly regarding access and retention issues. Paper presented at the 17th Annual Conference of the Society for Student Counselling in Southern Africa, Technikon Port Elizabeth.

Taylor, J.M. & Radford, E.J. 1986. Psychometric testing as an unfair labour practice. South African Journal of Psychology, 16 (3). 79-86.

Taylor, T.R. 1987. The future of cognitive assessment. Pretoria: Human Sciences Research Council.

Technikon Pretoria. 1998. Technikon Pretoria Student throughput statistics 1994-1997. Technikon Pretoria: Strategic Information and Planning, Intranet.

Thorndike, R.L. 1971. Concepts of culture-fairness. Journal of Educational Management, 8. 63-70.

Thorndike, R.L. 1986. The role of General Ability in Prediction. Journal of Vocational Behavior, 29. 332-339.

Van Zyl, M.A.; Terblanche, S. & Jacobs, G.J. 1992. Die profiel van 'n potensieel-suksesvolle eerstejaarstudent in Maatskaplike Werk: 'n Benadering tot keuring. Maatskaplike Werk, 28 (1). 33-43.

Van Aswegen, M. 1997. The standardisation of a learning potential battery for the selection of poorly qualified employees. Unpublished M Comm Thesis: University of Pretoria.

Walsh, W.B. & Betz, N.E. 1985. Tests and assessment. New Jersey: Prentice Hall.

Zaaiman, H. 1998. Selecting students for Mathematics and Science: The challenge facing higher education in South Africa. Unpublished Ph.D thesis: Vrije Universiteit, Amsterdam.

Zietsman, A.I. 1985. Prestasie in die matrikulasie-eksamen as voorspeller van sukses op universiteit. Unpublished research report: Rand Afrikaans University.