A statistical investigation into the admissions criteria for the Electrical- and Mechanical Engineering courses at the Peninsula Technikon

by

Desmond Reginald Booysen

Submitted in partial fulfillment of the requirements for the

Masters Degree in Commerce in the Department of Statistics

University of Western Cape

Prof. D. Kotze

Prof. J. Fresen

Date of submission: October 1995

Abstract

Concern was expressed by the School of Electrical- and Mechanical Engineering and Computer Data Processing (School of EMC) regarding the School's admissions criteria for the Electrical- and Mechanical Engineering courses. Admission to these courses was primarily based on candidates' matric performance. The results of an aptitude test were used as a secondary criterion in selecting "academically risky" students. This investigation focused mainly on the following issues:

1. *The use of cut-off points in matric results and an aptitude test as admissions criteria.*
2. *Identification of Matric subjects and/or aptitude test components that might assist in the selection of potentially successful candidates.*

This study reports on Chaid and discriminant analyses performed to identify the important predictors of successful study. From the results of these tests regression models were constructed that were used to determine expected pass rates in first semester subjects.

The results indicated that:

1. *The Swedish rating was the best predictor of successful first semester study. An admissions Swedish score of at least 35 is recommended.*
2. *The aptitude test total was found to be not significant in identifying potentially successful candidates. However, individual aptitude test components were significant in describing performance in specific first semester subjects.*

Acknowledgments

I would firstly like to thank my supervisors Professor D.Kotze and Professor J.Fresen for their guidance and support. A word of thanks are also extended to the HSRC in particular Professor D.J.Stoker and Mr D.Petersen for their invaluable assistance with the Chaid analyses. My heartfelt gratitude to Mrs M.Muller who had the onerous task of typing the report and Anwar Vahed for the computer expertise and support provided during the final stages of the work. A word of thanks to the University of the Western Cape for the financial support during my studies.

Sincere thanks to my wife Chesmaine and our children for their understanding, patience and encouragement. Without your support this work would never have come to fruition.

Finally, I would like to dedicate this work to my late Mother for the years of sacrifice, without which none of this would have been possible.

Table of Contents

Page

Page

Page

Page

List of Figures and Tables

Page

Page

Chapter 1

Introduction

1.1 Statement of the problem

During the selection of students for the Electrical- and Mechanical Engineering courses, heavy emphasis is placed on matric performance, in particular Mathematics and Physical Science. Borderline students are allowed to sit for an aptitude test.

The majority of students registered during the period under review came from socio-economically deprived backgrounds and inadequate schooling systems both of which influenced their performance in matric and/or the aptitude test detrimentally. A further complication was that candidates derived from a broad range of educational departments, each with differing standards.

1.2 Aim of the study

This study sets out to:

i) Determine which, if any, of the components of the aptitude test and/or matriculation results are important in predicting student success in Electrical and Mechanical Engineering at Peninsula Technikon.

ii) Establish the suitability of existing measuring instruments, namely, the aptitude test, Swedish rating and/or Matric aggregate in the selection process.

1.3 Literature review

As part of their search into establishing models according to which successful tertiary study could be described, Stoker et al conducted a broad literature study on the important parameters at play in the transition from school to tertiary institutions. This literature study has been used as the basis for my research.

According to Stoker et al a wide range of research that studied the effect of a variety of factors on successful tertiary study was done. Factors considered in these studies were amongst others biographical, school achievement, socio-economic, non-intellectual qualities (i.e. aptitude), study conditions and teaching strategies employed at tertiary institutions.

Although extensive research has been done in this field, Stoker could only identify four earlier studies in this field that were conducted at technikons. Of these studies, conducted between 1976 and 1982, three focused on the successful study of learner engineering technicians while the fourth focused on a range of technikon study fields.

Two important reasons why I undertook this study were:

1) Little research has been done on the problem of student selection at Technikons.
2) The typical cross-section of students attending the Peninsula Technikon may be different from the cross-section of students at those institutions where the research has been done.

1.4 Preview

Chapter 1 outlines the problem and the aims of the investigation.

In Chapter 2 the admissions criteria employed within the School concerning both Electrical- and Mechanical engineering students are outlined. It further highlights those factors considered important for successful study and which were readily available from the institution's records. Finally, the selection of the sample is discussed.

Chapter 3 presents an exploratory data analysis of the possible predictors of success discussed in Chapter 2. In particular, we considered the marginal distribution of these factors.

In Chapter 4 Chaid Analysis is briefly discussed followed by analyses of the possible predictors of successful first- and final semester study.

In Chapter 5 discriminant analysis is briefly discussed followed by discriminant analyses of first- and final semester predictor variables. Finally possible discriminants (predictors) for the identification of "academically risky" students are given. It further discusses the feasibility of using statistically developed models in the assignment of prospective candidates as either possible successes or failures.

Chapter 6 briefly outlines regression analysis. This is followed by an analysis to ascertain those predictor variables that are statistically significant in identifying success in individual first semester subjects. Expected pass-rates for individual first semester subjects, given different entry levels of proficiency in important predictors of success in these subjects, were then computed.

The final chapter outlines the findings of the study, attempts to highlight certain shortcomings and formulates recommendations.

Chapter 2

Admissions Criteria

2.1 Admissions Criteria

The admission requirements for the diplomas in Electrical and Mechanical Engineering, as set out in the brochure for the particular school, are as follows:

2.1.1 Electrical Engineering

The admissions criterion for Electrical and Mechanical Engineering courses was one of the following.

a) Senior Certificate requirements:

An E-symbol on the Higher Grade in both Mathematics and Physical Science or a D-symbol on the Standard Grade in both Mathematics and Physical Science.

OR

b) NTC 4-Certificate requirements:

A C-symbol in both Mathematics and Applied Science as well as the language requirement of the Joint Matriculation Board.

A further criterion was an admission mark, based on the Swedish Rating System (Peninsula Technikon 1991: 9), of not less than 29, or a minimum total stanine mark of 15 in the aptitude test for qualifying students.

2.1.1.1 Selection Procedure

Once the Matric- or NTC 4 Certificate criterion has been met, the students' results in the subjects listed in 2.1.1 (a) and (b) and the best of the remaining subjects, to a maximum of six subjects, were used to determine their admission status. Admission status refers to students' suitability for entry to the course. To this end the Swedish Rating System was employed which allows an admission mark to be calculated for each individual according to table 2.1 below:

Table 2.1: Subject performance rating using the Swedish rating system

Symbol	Rating		NTC3/4
	Higher Grade	Standard Grade	
A	8	6	6
B	7	5	5
C	6	4	4
D	5	3	3
E	4	2	2
F	3	1	1

The rating for each subject is determined according to the above scale while the ratings for Mathematics and Physical Science were doubled. Up to and including the year 1991, a candidate who obtained a Swedish rating of 29 and who satisfied the requirements for Mathematics, Physical Science/Applied Science in a/b above, would be regarded as suitable for the course. From January 1992, the Swedish rating was increased to 32.

A prospective student who did not satisfy the Swedish rating requirement, but met the requirements for Mathematics and Physical-/Applied Science, and provided further that the student gained a Swedish rating of not less than 28, he/she was allowed to sit for an aptitude test. The minimum mark (28) for admission to the aptitude test was increased to 31, effective as from January 1992. The admissions criteria for the period under review appear in table 2.2.

Table 2.2: Admissions criteria in terms of obtained Swedish rating levels

Swedish rating score	Admission
> 28	Accepted
28	Aptitude test
< 28	Not accepted

To summarise, the final admission of a student to the course depended on the number of places available, and on one or more of the following criteria:

a) Results in Mathematics, Physical Science/Applied Science;

b) Swedish rating;

c) Matric aggregate;

d) Results of aptitude test;

The school reserved the right to consider other factors in unusual circumstances. These factors are not relevant to the parameters of this thesis and will not be discussed here.

2.1.2 Mechanical Engineering

The initial selection criterion is:

a) **Senior Certificate requirements:**

A D-symbol on the Higher Grade **or** a C-symbol on the Standard Grade in both Mathematics and Physical Science.

OR

b) **NTC-3 Certificate requirements:**

A pass mark in Mathematics and Applied Science or equivalent subjects as well as a pass mark in both official languages.

Further criteria for admission were a Swedish rating of not less than 29 or a minimum total stanine mark of 15 in the aptitude test (See section 2.1.1.1).

2.1.2.1 Selection Procedure

The selection procedure was the same as for Electrical Engineering as stated in section 2.1.1.1.

2.2 Aptitude Test

2.2.1 Historical development of aptitude test:

The diversity of the South African Schooling system with its range of examining authorities and its inevitable differing standards and end-products posed serious challenges to the School of EMC in terms of admission to the engineering courses. In conjunction with the Student Advisory Bureau and with invaluable input from industry, the School of EMC set about finding ways to assist students from the academically deprived communities who may otherwise have been excluded from the Engineering courses if only the normal entry requirements were adhered to, an opportunity to gain access. To this end, several existing aptitude test batteries were obtained from the Human Sciences Research Council (HSRC). The Student Advisory Bureau in conjunction with the relevant departments in the School of EMC, then set about identifying suitable tests that could be of assistance in the selection of disadvantaged students. While the objective was to identify potential in these disadvantaged students to assist them to gain entry to the course, the team had to ensure that academic standards would not be compromised. The School chose four test batteries namely; Mental alertness, Technical reading comprehension, Mechanical comprehension and Gottschaldt test and suggested their immediate implementation with effect from 1987. These tests are still in use today. These test batteries were designed for a homogeneous student population of a particular race group and were therefore not suited for our diverse student population. Their use however constituted an additional selection measure in the absence of more suitable measures.

The intention was to establish appropriate norms for these tests at a later stage. Due to the unavailability of certain aptitude test components, only the Mental alertness test was administered during 1987, while the full scope of tests was implemented during 1988.

2.2.2 Aptitude test components

Owen et al (1988: 143), quoting Bingham (1937: 16), state that aptitude refers to a person's ability or future potential to acquire knowledge or skills, with training. The four test components administered to measure the aptitude of prospective students were:

Mental Alertness (High level battery):

According to Owen et al (1988: 140) the Mental alertness test measures the incumbent's verbal aspects of intelligence.

Technical Reading Comprehension:

The Technical reading comprehension test consists of five paragraphs chosen with the purpose of determining whether a testee with little or no formal schooling in technical subjects can comprehend articles of a technical nature (Van Zyl (1991:175)).

Mechanical Comprehension:

According to Owen et al (1988:231), the Mechanical comprehension test is designed to test the ability of students to appropriately apply the laws and principles of physics.

Gottschaldt:

According to Owen et al (1988:141) the Gottschaldt test provides a measure of the analytical capabilities required by the incumbent to locate tables hidden in complex diagrams.

Note: For purposes of later analysis, in particular Chaid analysis, first- and second language groups in both official languages were grouped into single categories, either as Afrikaans or English.

2.3.3 Aptitude test results

The factors considered were:

The raw and stanine scores obtained by individuals in each test;

Total raw scores and total stanine scores obtained by individuals.

2.3.4 Tertiary study record

The factors considered were:

Study course enrolled for namely: Electrical- or Mechanical Engineering;

Minimum duration of the study course;

Percentage obtained in the examination for each of the subjects enrolled for during the first-, second- and third semester of study;

Mean percentage obtained for all subjects enrolled for during a specific semester;

First semester achievement and the time that elapsed until first semester success was achieved (first semester success refers to a student passing all seven prescribed first semester courses);

Final semester achievement and the time that elapsed until success was achieved (final semester success refers to completion of only the theoretical component of the full course), or in the case of unsuccessful candidates, the number of semesters the student has been registered for the particular course.

2.3.2 School study record

The factors considered here were:

Type of matriculation or equivalent qualification obtained. For example full matriculation exemption, ordinary-, mature age-, foreigners- and immigrants conditional exemption, senior certificate without exemption, NTC3/4 with matric languages, Practical Matric and a category for those types not specifically mentioned;

The examining body from which the qualification could be obtained namely Joint Matriculation Board (JMB), Cape Education Department (CED), Transvaal Education Department (TED), Natal Education Department (NED), National Education: RSA (RSA), Departments of Education and Culture: House of Assembly (HOA), House of Representatives (HOR) and House of Delegates (HOD), Department of Education and Training (DET), National Education: SWA (SWA), Transkei Education (TE) and a class for those examining bodies not specified (Other);

Specific subjects offered in the final examination, the level at which these were written and the students' level of achievement in the subjects: Afrikaans First Language, Afrikaans Second Language, English First Language, English Second Language, Black Languages, History, Geography, Physical Science, Biology, Mathematics, Biblical Studies, Accounting, Business Economics, Economics, Woodwork, Electrician's Work, Fitting and Turning, Metalwork, Technical Drawings and Agricultural Science and

Aggregate obtained in the final examination.

Admission to the aptitude test depended on the Swedish rating and the availability of places. The marks scored by students in the individual components of the aptitude test, were transformed into stanines. A total stanine mark for each student was then calculated by summing the stanine values students obtained in each component of the aptitude test. These total stanine values were then used to determine prospective students' eligibility for the course. Those candidates who obtained a total stanine mark of at least fifteen (15), irrespective of individual test marks, were then shortlisted for possible admission to the course. Depending on the number of places available, those candidates with the highest total stanine values were admitted.

2.3 Factors considered in this study

The factors considered in this study were based on those identified by Stoker et al (1985: 9 - 10) as being relevant to the problem of student selection and their availability from student records. These factors can be categorised as follows:

Personal particulars

School study record

Aptitude test results

Tertiary study record

2.3.1 Personal particulars of each student

The factors considered were:

Age;

Gender;

Home language: This was categorised into Afrikaans, English, Afrikaans/English and a Black language, and

Activity before entering Technikon: Seven distinct categories of applicant were identified namely University student, Technikon student, Technical College student, National service, Labour force, Matriculant and a category termed 'Other', to cover activities not specifically mentioned.

Note: Students were considered to be registered for a particular course even if he/she was only registered for a single subject which they had failed during a previous semester.

2.4 The sample

The population was defined as all Electrical and Mechanical Engineering students ever enrolled at the Peninsula Technikon. The sample selected consisted of all first time Electrical and Mechanical registrations at this Technikon for 1988 and 1989. This is a convenience rather than a random sample, but it was a logical choice since full scale aptitude results (an important factor in the study) were only obtained for the first time in 1988. The second reason was that records of the students' progress in the theoretical component of the course were readily available.

Student records for the period under review, namely 1988 to 1991, were obtained from the Technikon's student files. Matric results reflected on the student files were verified against copies of their matriculation- or other appropriate certificates. Students for whom the required information as outlined in section 2.3 was available, were included in the study. Given the criteria outlined in section 2.3, exclusion from the sample was based on the following:

a) Insufficient or no matric data available;

b) Insufficient or no tertiary study record available.

Based on these two exclusion criteria, the final sample consisted of five hundred and eighty-four of the 1988 and 1989 first-time registrations.

This sample consisted of 282 students registered in 1988 of which 150 were Electrical Engineering students and 132 Mechanical Engineering students. The remaining 302 students that registered in 1989 consisted of 180 Electrical Engineering and 122 Mechanical Engineering students.

Chapter 3

Exploratory Analysis of Pass Rates

3.1 Introduction

An exploratory data analysis was performed on all the factors that we believed might play an important role in explaining the variation of students' performance at the Technikon. The aim of the analysis was to assist in highlighting the most important features of our set of data and possibly point the way to an appropriate analysis path.

3.2 Summary

1. The very low pass rates call to question the preparedness of our students for Electrical- and Mechanical Engineering studies. The overall pass rate was 46% for the first semester studies and 12,3% for the full course.

2. A further disturbing factor was that only 25% of the student body passed the first semester courses at the first attempt, while a mere 5,7% of the students passed the full course at the first attempt.

3. The first semester- and full course pass rates improved with improved Swedish rating or Matric aggregate.

4. The results supported the School's decision to increase the Swedish rating admission requirement to above 30, as those students with scores below 30 have very little chance of success.

5. The pass rates of students with Swedish rating scores above 32 improved significantly. Even at first semester level, pass rates in excess of 60% was only attained at Swedish rating levels of at least 35.

6. Students with labour force experience seemed to perform better than any other group. Their performances, in both the first semester courses and the full course, were to a large extent accomplished at the first attempt.

7. A disturbing aspect was the poor performance of students directly from school. This is particularly troublesome since this group constituted the bulk of the student body.

8. The poor performance of students with a Practical Matric, indicated that the continued use of this qualification as an entrance qualification for the Electrical - and Mechanical Engineering courses, should be closely monitored with a view to a possible review.

9. Of the three major contributors to our student body, namely the HOR, HOA and DET, those with a DET qualification performed the best. The worst performers were students from the HOR group. It might be in the interest of the Technikon if the underlying reasons for the overall poor performance of the students, in particular the HOR group, are investigated at some stage.

10. Afrikaans speaking students seemed to be at a disadvantage. Their pass rates were substantially lower than those of any other group.

11. No definite conclusion regarding the aptitude aggregate could be made, possibly because only students in a very narrow performance range were available for the analysis of the aptitude test performance.

3.3 Distributions of personal characteristics and study record

In this section we analyse the marginal distributions of each of the factors considered in the study. We also consider the pass rates for these factors and give a short discussion of the results. Finally, a brief overview of student performance in individual Matric and Technikon subjects are given.

3.3.1 Distribution of Swedish rating

The distribution of students according to Swedish ratings and their corresponding pass-rates amongst groups for first semester and full course performance are presented in tables 3.1 and 3.2. Full course performance refers to the successful completion of the theoretical component of the course as a whole. Although the admissions cut-off point for the Swedish rating was 29, these tables indicate that a considerable number of students had Swedish ratings less than the required minimum. This may be because Matric aggregate, Mathematics and Physical Science performance, and not only Swedish rating marks, were used as selection criteria. A significant number of students were also admitted on the basis of their aptitude test performance.

3.3.1.1 First semester performance

Table 3.1: Student performance in first semester studies according to the Swedish rating

Swedish Rating	Frequency	Percentage	Duration of Study for First Semester Success (Semesters)			Total Number of Successful Students	Percentage Successful Students per Performance Group
			1	2	>2		
10-14	15	3,1	1	0	2	3	20,0
15-19	55	11,5	9	3	8	20	36,4
20-24	74	15,4	9	4	6	19	25,7
25-29	115	24,0	17	11	22	50	43,5
30-34	108	22,5	29	12	16	57	52,8
35-39	58	12,1	28	7	5	40	69,0
40-44	41	8,6	25	6	1	32	78,0
Above 44	13	2.7	9	2	1	12	92,3
Total	479	100	127	45	61	233	48,6

Note: i) Missing data in the compilation of this table and table 3.2 occured because sufficient detail was not always available to accurately compute Swedish rating scores for all students.

ii) The sample size for students in the full course (refer Table 3.2) is larger than that for first semester students (Table 3.1). This apparent contradiction is due to the fact that some students who transferred from other tertiary institutions were credited for first semester courses succesfully completed at those institutions and the details of these results were not always readily available.

While 54% of all the students had Swedish ratings of less than 30, this group produced only 39% of the successful students. Those students who had Swedish ratings of at least 30, constituting 61% of the successful students, had an average pass-rate of 64%.

The last column of table 3.1, representing the conditional pass rates at given Swedish rating levels, indicated that, except for an unexplained dip in the pass rate at the third group, there was a progressive improvement in the pass-rate with improved Swedish rating mark. Only 20% of the students in the 10-14 Swedish rating group were successful. The pass rate improved dramatically with increased Swedish rating to approximately 92% for the students in the above-44 group. Around a Swedish rating of approximately 32 (the revised cut-off point) the pass rate was 52,8 %.

There was also a general improvement in the first-time pass rate with improved Swedish ratings. Furthermore, 127 of the 233 successful students passed their first semester courses at the first attempt. This gives a first-time pass rate of 27%. (First-time pass rate is defined as the percentage of students who completed their studies at the first attempt.)

From these results there seems to be a correlation between Matric performance as measured by the Swedish rating, and first semester success. Low Swedish ratings seem to point to low pass rates and high Swedish ratings to high pass rates.

There is also strong indications that restricting admission to candidates with Swedish ratings of at least 32 would greatly improve the probability of success. This step would however exclude a considerable number of prospective students. The Technikon Management may therefore wish to consider the merits and demerits of such a decision on other criteria.

3.3.1.2 Final performance

Table 3.2: Student performance in full course according to the Swedish rating

Swedish Rating	Frequency	Percentage	Duration of Study for Full Course Success (Semesters)			Total Number of Successful Students	First -time pass rate (Percentage)	Percentage Pass-rate per Performance Group
			3	4	>4			
10-14	15	3,0	0	0	0	0	0,0	0,0
15-19	57	11,3	1	2	2	5	1,8	8,8
20-24	77	15,2	1	1	1	3	1,3	3,9
25-29	123	24,3	6	3	4	13	4,9	10,6
30-34	114	22,5	3	6	3	12	2,6	10,5
35-39	66	13,0	7	3	3	13	10,6	19,7
40-44	41	8,1	7	5	1	13	17,1	31,7
Above 44	13	2,6	5	1	0	6	38,5	46,2
Total	506	100,0	30	21	14	65	5,9	12,8

Similar patterns as those that evolved for first semester success were also observed here. The percentage pass rates initially increased, dipping at the third category, whereafter it again increased steadily. No explanation could be found for the dip in the pass rate at that particular interval. With minor deviations, the general trend was that the pass rate improved with improved Swedish rating (See figure 3.3.). Pass rates at the lower end of the Swedish rating scale ranged from 0% to approximately 11% at an average Swedish rating of 32. The pass rate improved to 46% at the above-44 Swedish rating level.

From tables 3.1 and 3.2 it is clear that for certain Swedish rating levels, the overall pass rate at both first- and final semester level was extremely low. The very low pass rate attained by students in most of the categories was also disturbing. While 53% of all candidates had Swedish ratings of less than 30, this group only constituted 32% of the successful students. Students in the categories with Swedish ratings below 35 had pass rates below the overall pass rate of 12,8%. Above the Swedish rating of 35, pass rates were nearly one and a half to four times higher than the overall pass rate.

Only thirty of the students passed at their first attempt. This translates into approximately 6% of the total sample size used in this investigation. The poor performance of the majority of the students, both in terms of first-time pass rates as well as conditional pass rates, would indicate that our selection criteria, in particular the Swedish rating cut-off point, could be inadequate and that serious attention would have to be given to this aspect. Figure 3.3 for example suggests that in order to achieve a pass rate of say 50% in the first semester, the Swedish rating cut-off point should be approximately 32. If, however, the concern was for achieving a full course pass rate of, say, 50%, a Swedish rating level above 52 would be required.

Figure 3.3: Graph of student performance in first semester and full course according to the Swedish rating.

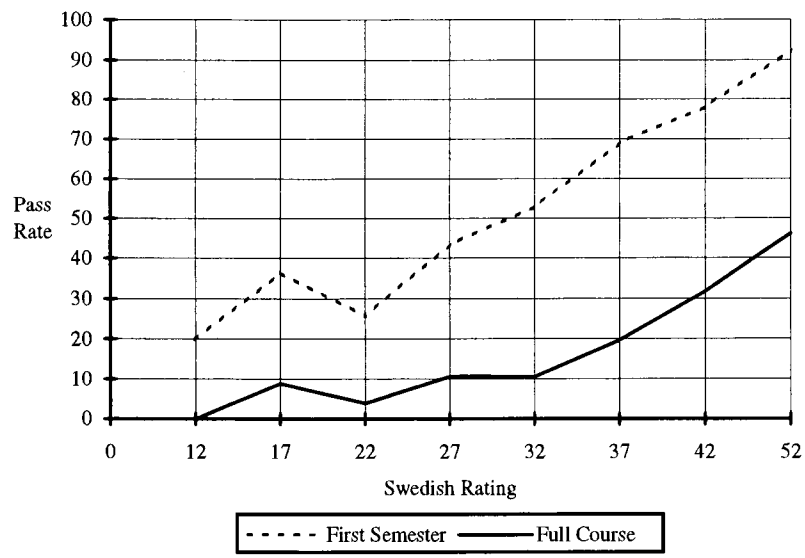


Figure 3.3 also demonstrates that, vastly different pass rates are attained at the first semester- and the full course. The Swedish rating could possibly be successfully manipulated within reasonable limits, say, Swedish rating levels of between 30 and 40, to significantly improve the pass rate at first semester level. Attaining acceptable pass rates in the full course through manipulation of the Swedish rating would however be more difficult. This is so since pass rates for the full course are significantly lower than first semester pass rates at corresponding Swedish rating levels, and full course pass rates only give reasonable pass rates at very high Swedish rating levels.

3.3.2 Distribution of matric aggregate

The distribution of students according to matric aggregate together with pass-rates amongst groups for first- and final semester performance are given in tables 3.4 and 3.5 respectively. Missing values in both tables 3.4 and 3.5 and can be ascribed to insufficient detail of matric results available as well as non-availability of detail of student performance in either the first semester (mainly due to transfers from other institutions) or final semester (due to e.g. drop-outs, transfers to other institutions, incomplete coursework,etc).

3.3.2.1 First Semester Performance

Table 3.4: Student performance in first semester studies according to overall matric performance

Matric Aggregate (%)	Frequency	%	Duration for First Semester Success (Semesters)			Total Number of Successful Students	First time pass rate (Percentage)	% Successes per Matric Aggregate
			1	2	>2			
33 - 49	220	42,5	33	15	33	81	15,0	36,8
50 - 59	213	41,1	62	20	30	112	29,1	52,6
60 - 79	85	16,5	43	10	7	60	50,6	70,6
Total	518	100,0	138	45	70	253	26,6	48,8

Because of the low sample size of the 33 - 39% category, this category was combined with the 40 - 49% category. Similarly the 60 - 69% and 70 - 79% categories were combined so that meaningful discussion of the conditional pass rates would be possible.

There was a distinct improvement in the pass rate from 36,8% to 70,6% in the first semester with an increase in Matric aggregate. Students in the 60 - 79% group had a pass rate of nearly double that of the 33 - 49% group. Furthermore, we observed that the pass rate of the 33 - 49% group was well below the average pass rate. Since this group constituted nearly 43% of the student body, their poor performance had a significant impact on the overall performance. A further important observation was that the percentage of students who successfully completed their studies within one semester, increased significantly with improved aggregate. For example, the first-time pass rate was only 15% at the 33 - 49% aggregate level, but it increased to 50,6% at the 60 - 79% aggregate level. Students at the upper end of the aggregate scale were therefore approximately three times as successful at the first attempt than those students at the lower end of the aggregate scale. In conclusion, low conditional pass rates and low first-time pass rates corresponded with low matric aggregates, while high conditional pass rates and high first-time pass rates corresponded with high matric aggregates. This was strong evidence that performance in the first semester is correlated with matric performance as measured by their matric aggregate.

3.3.2.2 Final Performance

The categories in table 3.5 were combined in a similar manner as those in table 3.4 in order to facilitate more meaningful discussion of the pass rate distribution.

Table 3.5: Student performance in full course according to overall matric performance

Matric Aggregate (%)	Frequency	%	Duration of Study for Full Course Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Percentage Successes per Aggregate Group
			3	4	>4			
33 - 49	220	42,5	5	6	5	16	2,3	7,3
50 - 59	213	41,1	14	10	8	32	6,6	14,8
60 - 79	85	16,5	11	5	2	18	12,9	21,2
Total	518	100,0	30	21	15	66	5,8	12,6

There was a significant increase in the conditional pass rate from 7,3% to 21,2% with increased matric aggregate. Students in the 60 - 79% matric aggregate category were more than three times as successful as those in the 33 - 49% matric aggregate category. We further observed that the pass rate for the 33 - 49% group was well below the overall pass rate. Furthermore, the first-time pass rate of the students in the 33 - 49% matric aggregate category was only 2,3%. The corresponding figures for the next two matric aggregate categories, although disappointingly low at 6,6% and 12,9% respectively, were a significant improvement on the first-time pass rates obtained by the students at the bottom of the matric aggregate scale. There was therefore a significant increase in the first-time pass rate with improved matric aggregate. These results suggest that in order to improve the pass rate to a more acceptable level both for the first semester and full course, admission level aggregates should be increased.

However, the low pass rates, even at high aggregate levels, indicate that this measure alone would not guarantee satisfactory levels of success. It would however go a long way in raising success levels, particularly at first semester level. Again, this is strong evidence that overall student performance is correlated to matric performance as measured by their matric aggregate.

3.3.3 Distribution of Activity before entering technikon

The majority of students had either been matric pupils (48.3%) or were previously registered at another Technikon (28.3%) during the period immediately before registering at Peninsula Technikon. Tables 3.6 and 3.7 depict the distribution of students according to activity before entering Peninsula Technikon, and the pass rates amongst activity groups for first semester- and final performance respectively.

3.3.3.1 First Semester Performance

Table 3.6: Student performance in first semester studies according to activity before their first enrollment at Peninsula Technikon

Activity before Technikon entry	Frequency	%	Duration for first semester success (Semesters)			Total number of successful students	First-time pass rate (Percentage)	Percentage successes per activity group
			1	2	>2			
Labour Force	29	5,0	18	2	2	22	62,1	75,9
Technikon student	165	28,3	35	23	37	95	21,2	57,6
Other	58	9,9	25	3	4	32	43,1	55,2
Technical College	35	6,0	15	2	2	19	42,9	54,3
Matric Pupil	282	48,3	51	18	28	97	18,1	34,4
University student	15	2,6	3	0	0	3	20,0	20,0
Total	584	100,0	147	48	73	268	25,2	45,9

Students with labour force experience, with a pass rate of 75,9%, out-performed all other students, followed by three categories of students, namely Technikon students, Technical College students and those termed Other with pass rates of around 55%. The students with the worst results were the Matric students with a pass rate of 34,4% and University students with a pass rate of 20%. It is disturbing to note that students entering Technikon directly from school performed worse than all but one other group, especially as they comprised approximately 48% of the student body.

The first-time pass rates varied from 18,1% to 62,1%. Matric students had the worst first-time pass rate. The largest discrepancies between the conditional pass rates (column 9) and first-time pass rates (column 8) were observed for students in the two largest groups, namely Technikon - and Matric pupil groups. It is therefore not surprising that the overall first-time pass rate was only 25,2%. This phenomenon is undesirable half of the successful students took longer than one semester to complete their first semester courses, and thus competed for space (a scarce commodity) with prospective students during the following academic periods.

3.3.3.2 Final Performance

Table 3.7: Student performance in full course according to activity before their first enrollment at Peninsula Technikon

Activity Before Technikon Entry	Frequency	%	Duration for Full Course Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Percentage Successes per Activity Group
			3	4	>4			
Labour Force	29	5,0	6	2	0	8	20,7	27,6
Other	58	9,8	7	3	0	10	12,1	17,2
Technical College	35	6,0	2	2	2	6	5,7	17,1
Technikon Student	165	28,3	6	10	9	25	3,6	15,2
Matric Pupil	282	48,3	12	6	5	23	4,3	8,2
University Student	15	2,6	0	0	0	0	0,0	0,0
Total	584	100,0	33	23	16	72	5,7	12,3

Students from the labour force again out-performed all other students. This group had a pass rate of more than twice the overall pass rate, and far exceeded all the other conditional pass rates. Only the University students had a worse pass rate than the two major groups, namely Matric students and Technikon students. Technikon students, (with a pass rate of 15,1%), were nearly twice as successful as Matric students, (with a pass rate of 8,2%). Since these two groups constituted the majority of our students, their very low pass rates had a severe effect on the overall pass rate.

Approximately 3,6% of the Technikon students and 4,3% of the Matric pupils managed to complete the full course at the first attempt. The first-time pass rate of the Technikon students comprised only 25% of the successful students in this group. This means that approximately 75% of the students in this group took more than three semesters to complete the full course. The first-time pass rate of the Matric students indicated that approximately 50% of the students in this group took more than three semesters to successfully complete the full course.

The best first-time pass rate, which was attained by the labour force group, was nearly five times as high as those attained by the two major groups, namely Matric- and Technikon students. Furthermore, the first-time pass rate for the labour force group showed that only 25% of the successful students from this group needed more than three semesters to complete the full course. The poor results posted by the two major groups, namely Matric- and Technikon students, have serious academic and financial implications for the School and hence the Technikon as institution. It is therefore advisable that close attention be given to these groups.

3.3.4 Distribution of Pre-Technikon qualification type

Distribution of students with respect to pre-technikon qualification type in terms of pass rates amongst the different classes, are reflected in tables 3.8 and 3.9. From table 3.8 it is clear that while 47,1% of students had a Senior Certificate without matriculation exemption, approximately 36% of the students had some form of exemption. A significant percentage (21,9%) were in possession of a technical pre-technikon qualification.

3.3.4.1 First Semester Performance

Table 3.8: Student performance in first semester studies according to matric status

Matric Type	Frequency	%	Duration for First Semester Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Percentage Successes per Matric Type
			1	2	>2			
Full exemption	53	9,1	20	6	4	30	37,7	56,6
Ordinary Conditional exemption	90	15,4	26	6	12	44	28,9	48,9
Other	38	6,5	10	3	5	18	26,3	47,4
Senior Certificate without exemption	275	47,1	65	25	32	122	23,6	44,4
NTC3 including matric languages	41	7,0	11	1	6	18	26,8	43,9
Matric Practical	87	14,9	16	6	12	34	18,4	39,1
Total	584	100	148	47	71	266	25,3	45,6

The majority of students entered the Technikon with a Senior Certificate without exemption. The next largest group were those students having a Conditional exemption followed by students with a Practical Matric qualification. The Full exemption group had both the highest conditional pass rate and the highest first-time pass. Those with a Practical Matric qualification attained the poorest results, in terms of both the conditional pass rates as well as first-time pass rates. Students with a Senior Certificate without exemption, had a pass rate of 44,4% and a first-time pass rate of only 23,6%. Although students from this group would therefore have a 44% chance of achieving success, only about one in four students would achieve success at the first attempt.

3.3.4.2 Final Performance

Table 3.9: Student performance in full course according to matric status

Matric Type	Frequency	%	Duration for Full Course Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Percentage Successes per Matric Type
			3	4	>4			
Full exemption	53	9,1	5	5	1	11	9,4	20,8
NTC3 including matric languages	41	7,0	2	3	2	7	4,9	17,1
Senior Certificate without exemption	275	47,1	16	9	11	36	5,8	13,1
Ordinary Conditional exemption	90	15,4	6	4	0	10	6,7	11,1
Other	38	6,5	1	1	2	4	2,6	10,5
Practical Matric	87	14,9	3	1	0	4	3,4	4,6
Total	584	100	33	23	16	72	5,7	12,3

The pass rates for all groups were generally low, ranging from 4,6% - 20,8%. The Full exemption group had the highest percentage of successful students (20,8%) as well as the highest percentage of students who successfully completed the full course at the first attempt (9,4%). While 13,1% of the students in the Senior Certificate without exemption group were successful, only 5,8% of the students in this group managed to pass the full course at the first attempt. This means that more than half of the students in this group took more than the minimum of three semesters to successfully complete the full course. The Practical Matric group not only had the dubious honour of having the worst pass rate (4,6%), but the overall pass rate was nearly three times as high as the pass rate of this group. An interesting phenomenon was the emergence of the NTC3 students as the group with the second highest pass rate. Serious reservations regarding the usefulness of a Practical Matric as an admissions qualification for Electrical and Mechanical Engineering studies, were raised by the poor performance of the students in this group. Their poor performance is of particular concern because, as the group with the third highest number of students, (14,9%), they constituted a significant component of the student body.

3.3.5 Distribution of Examining body

The distribution of students according to the examining body, together with pass rates for each, are given in tables 3.10 and 3.11. These tables are given in order of decreasing pass rate for both first semester- and final performance .

3.3.5.1 First Semester Performance

Table 3.10: Student performance in first semester according to examining body

Examining Body	Frequency	%	Duration for First-Semester Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Percentage Successes per Examining Body
			1	2	>2			
Other	45	7,9	21	6	4	31	46,7	68,9
Transkei Education	19	3,3	3	2	5	10	15,8	52,6
Joint Matriculation Board	12	2,1	5	0	1	6	41,7	50,0
Education & Training	71	12,4	22	6	7	35	31,0	49,3
Education & Culture: Assembly	29	5,1	10	0	3	13	34,5	44,8
Education & Culture: Representatives	351	61,5	75	29	49	153	21,4	43,6
National Education: RSA	12	2,1	2	2	1	5	16,7	41,7
Cape Education	16	2,8	4	1	1	6	25,0	37,5
Education & Culture: Delegates	16	2,8	3	1	0	4	18,8	25,0
Total	571	100,0	145	47	71	263	25,4	46,1

It is interesting to note that apart from the various local examining bodies, a few foreign examining bodies were also represented. The three major examining bodies represented were the Department of Education & Culture: House of Representatives, (HOR), with approximately 62% of the student body; the Department of Education and Training (DET) with 12,4% and the Department of Education & Culture: House of Assembly (HOA) with 5,1%. Those examining bodies with low sample sizes

namely TED, NED and SWA have been included in the group termed 'Other' in order to make discussion of the pass rates more meaningful.

The highest pass rate (68,9%) was attained by the 'Other' group. This was followed by the TE, JMB and DET groups with pass rates of approximately 50%. Students from the HOR which accounted for 61% of the student population had the sixth best pass rate at 43,6%. The worst results were attained by students from the HOD group. The pass rates of the three major groups ranged between 43,6% and 49,3%. The HOR students take on the average longer than their DET and HOA counterparts to complete their first semester courses. For example, only 21,4% of HOR students completed the first semester courses at the first attempt, while 31,0% of the DET students and 34,5% of HOA students were able to accomplish this. Since these three groups accounted for approximately 80% of the student population, their mediocre performance holds serious financial and academic implications for the School.

3.3.5.2 Final Performance

The JMB students performed the best, with a pass rate of 33,3%, followed by TE students (26,3%). The HOR students, who accounted for the majority of our student population, had a pass rate of 9,1%. This was in fact the worst pass rate attained by students from any of the three major examining bodies. The best results were attained by DET students (21,1%). The DET students were in fact more than twice as successful as students from the HOR group.

Only 5,6% of the DET students and 4,6% of the students in the HOR group were successful at the first attempt. Only 3,4% of the HOA students passed the full course at the first attempt. The dismal first-time pass rates and conditional pass rates of the three major groups indicated that few of the students from these groups were adequately prepared for their studies.

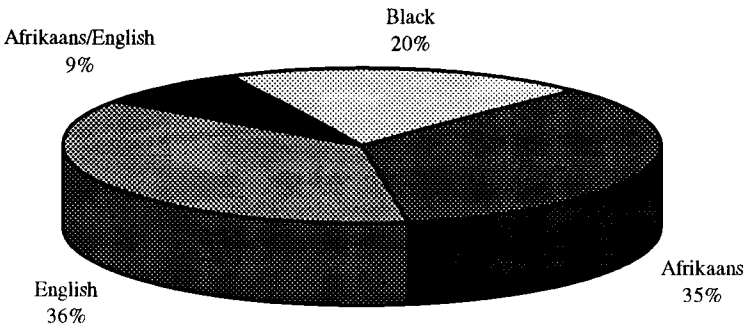
Table 3.11: Student performance in full course according to examining body

Examining Body	Frequency	%	Duration for Full Course Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Percentage Successes per Examining Body
			3	5	>4			
Joint Matriculation Board	12	2,1	3	1	0	4	25,0	33,3
Transkei Education	19	3,3	2	1	2	5	10,5	26,3
Education & Training	71	12,4	4	6	5	15	5,6	21,1
Other	45	7,9	5	4	0	9	11,1	20,0
National Education: RSA	12	2,1	1	1	0	2	8,3	16,7
Education & Culture: Assembly	29	5,1	1	0	2	3	3,4	10,3
Education & Culture: Representatives	351	61,5	16	9	7	32	4,6	9,1
Education & Culture: Delegates	16	2,8	0	1	0	1	0,0	6,3
Cape Education	16	2,8	0	0	0	0	0,0	0
Total	571	100,0	32	23	16	71	5,6	12,4

3.3.6 Distribution of Home language

The home language of approximately 81% of the students in the sample was either Afrikaans or English or both, while the rest spoke one of the Black languages. Forty five comma six percent (45,6%) of the students used English only or both English and Afrikaans as home language. A significant number of students (35%) were Afrikaans- speaking only. Figure 3.12 represents a percentage distribution of home language spoken by students. Since the medium of instruction at the Technikon is primarily English, it is important to note that approximately 55% of the students received tuition in a language other than their mother-tongue. This has obvious drawbacks for the learning process.

Figure 3.12: Composition of sample according to home language



3.3.6.1 First Semester Performance

Analysis of success rate in terms of home language is given in tables 3.13 and 3.14. These tables reflect the marginal distribution of home language and pass-rates amongst language groups for first semester- and full course performance respectively. The tables are given in order of decreasing pass rates.

Table 3.13: Student performance in first semester studies according to home language

Home Language	Frequency	%	Duration for first semester success (Semesters)			Total number of successful students	First-time pass rate (Percentage)	Percentage of successes per language group
			1	2	>2			
Black	114	19,5	37	10	15	62	32,5	54,4
English	216	37,0	63	19	25	107	29,2	49,5
Afrikaans/ English	50	8,6	14	1	7	22	28,0	44,0
Afrikaans	204	34,9	34	17	25	76	16,7	37,3
Total	584	100,0	148	47	72	267	25,3	45,7

The highest pass rate in first semester studies, (54,4%), was attained by the group using a Black language as mother tongue. The worst results were attained by the Afrikaans speaking group whose pass rate was significantly lower than those of the other three groups. While the other three groups had first-time pass rates of around 30%, the first-time pass rate for the Afrikaans speaking group was only 16,7%.

Approximately 46% of all the students completed their first semester course, but only about 25% of the students managed to do so in the minimum period of one semester.

3.3.6.2 Final Performance

Table 3.14: Student performance in full course according to home language

Home Language	Frequency	%	Duration for full course success (Semesters)			Total number of successful students	First-time pass rate (Percentage)	Percentage successes per language group
			3	4	>4			
Black	114	19,5	8	9	6	23	34,8	20,2
Afrikaans/ English	50	8,6	5	1	2	8	62,5	16,0
English	216	37,0	15	7	5	27	55,6	12,5
Afrikaans	204	34,9	5	6	3	14	35,7	6,9
Total	584	100,0	33	23	16	72	45,8	12,3

The pass rate per language group was alarmingly low, ranging from 6,9% to 20,2%. It is interesting to note that not only did the Black language group perform the best but that their pass rate was nearly three times that of the Afrikaans speaking group, who performed the worst. The poor performance of the Afrikaans speaking group was cause for concern since this group constituted 35% of the student population.

A further matter of concern was that only 5,7% of all the students passed in the minimum period of three semesters. Although the overall performance of the Black language group surpassed that of all other groups, its first-time pass rate was worse than that of the bilingual group, and approximately equal to that of the English speaking group.

The overall poor performance of all groups have serious financial, space and other implications for the institution and should therefore be given serious consideration.

3.3.7 Distribution of Aptitude test results

The marginal distribution of the aptitude test components as well as pass rates in first semester studies and the full course are given in tables 3.15 to 3.17. An analysis of performance in the different test components is reflected in table 3.15.

Table 3.15: Student performance in individual aptitude test components and overall

Variable	Frequency	Mean	Standard Deviation	Coefficient of Variation	Mode	Range	Median	Q_3-Q_1
Mental Alertness	358	45,16	15,69	34,74	31	83	44	24
Reading Comprehension	358	67,11	20,83	31,04	93	80	73	20
Mechanical Comprehension	358	45,59	12,89	28,23	50	81	45	14
Gottshaldt	358	33,99	17,13	50,41	22	93	29	22
Total	358	48,35	12,28	25,39	48	79	48	16

The student performance ranged from 34% for the Gottschaldt test to 67% for the Reading Comprehension test. Table 3.15 further indicated that the largest variations in student performance occurred in the Gottschaldt test. This points to poor analytical ability on the part of the students. The results further indicated that students coped well with the comprehension of technical articles.

3.3.7.1 First Semester Performance

Because of low sample size, the first two as well as the last two aptitude test total categories were combined in order to facilitate more meaningful discussion of pass rates. This was done for student performance in both the first semester studies, (table 3.16), and the full course (table 3.17).

Table 3.16: Student performance in first semester studies according to aptitude test performance

Aptitude test total (%)	Frequency	Percentage	Duration of Study for First Semester Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Percentage Successes per Aggregate Group
			1	2	>2			
Below 31	22	6,2	6	1	5	12	27,3	55,0
31 - 40	61	17,0	10	5	6	21	16,4	34,4
41 - 50	132	36,9	30	11	20	61	22,7	50,0
51 - 60	85	23,7	25	7	13	45	29,4	52,9
Above 60	58	16,2	24	5	0	29	41,4	50,0
Total	358	100,0	95	29	44	168	26,5	46,9

No definite pattern was observed when pass rates within the different categories were considered. Except for the inexplicable drop in the pass rate from 55% to 34,4% between the first and second categories and the subsequent increase to 50,0% at the third category, the pass rate remained relatively stable between values of 50% to 55%. It is anomalous that the group that had the worst aptitude test results attained the best pass rate (also refer to figure 3.18). The first-time pass rate, from the third category onward, showed an increase in pass rate with an increase in aptitude test total. In all aptitude test total categories, except the last one, there were large discrepancies in the conditional pass rates (column 9) and the first-time pass rates (column 8). The first-time pass rates in the first three categories were in fact less than half the corresponding conditional pass rates, while it was slightly more than 55% in the fourth category and in excess of 80% for the final category. This means that the majority of students in the first three categories, (those with aptitude test total up to 50), took longer than the minimum time to complete the first semester courses. Above an aptitude test total of 50, we observe a progressive improvement in the number of students who successfully completed the first semester courses at the first attempt. The usefulness of the aptitude test total in predicting first semester success is therefore questionable.

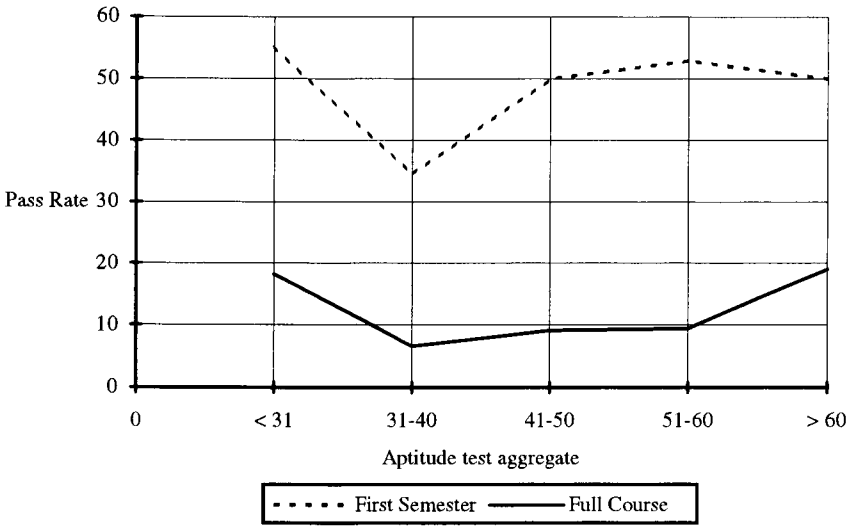
3.3.7.2 Final Performance

Table 3.17: Student performance in full course according to aptitude test performance

Aptitude test total (%)	Frequency	Percentage	Duration of Study For Full Course Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Percentage Successes per Aggregate Group
			3	4	>4			
Below 31	22	6,2	2	0	2	4	9,1	18,2
31 - 40	61	17,0	0	1	3	4	0,0	6,6
41 - 50	132	36,9	5	4	3	12	3,8	9,1
51 - 60	85	23,7	4	2	2	8	4,7	9,4
Above 60	58	16,2	6	2	3	11	10,3	19,0
Total	358	100,0	17	9	13	39	4,7	10,9

Students in most of the categories had a first-time pass rate of less than 10%. The overall first time pass rate accounted for approximately 5% of the students. It is interesting to note that the highest pass rates within groups occurred in the tails of the distribution with 18,2% and 19% for the first and last groups respectively. These pass rates were approximately twice the average pass rate and nearly three times as high as that of the worst group. From the second category onward, there was a progressive, although slow, increase in both the first-time pass rates and the conditional pass rates. There was very little distinction between the pass rates of students in the third and fourth categories. As shown in figure 3.18, two very divergent groups, namely those with the worst and those with the best aptitude test results, attained the best pass rates. I could find no explanation for this phenomenon. It is also clear from figure 3.18 that whatever the relationship was between the pass rates and the aptitude test performance, that the strength of this relationship decreased with time. This is reflected in the decrease in the pass rates from the first semester level to the full course level, at corresponding aptitude test total levels.

Figure 3.18: Graph of student performance in first semester and full course according to aptitude test performance.



3.3.8 Distribution of gender

Tables 3.19 and 3.20 below depict the distribution of students according to gender, together with pass rates amongst gender groups, for first - and final semester performance respectively.

3.3.8.1 First Semester Performance

Table 3.19: First semester performance according to gender

Gender	Frequency	Percentage	Duration for first semester success (semesters)			Total number of successful students	First-time pass rate (Percentage)	Pass Rate as a percentage of successful students per gender group
			1	2	>2			
Female	17	2,9	4	1	3	8	23,5	47,1
Male	567	97,1	144	48	68	260	25,4	45,9
Total	584	100,0	148	49	71	268	25,3	46,0

Since the number of female students was very low, results obtained from gender comparisons cannot be considered as very significant.

There was very little difference, in both the conditional pass rates and the first-time pass rates, between the genders. While 23,5% of the females completed their first semester studies at the first attempt, the corresponding figure for males was 25,4%. However slightly more females than males eventually passed the first semester course (47,1% versus 45,9%).

3.3.8.2 Final Performance

A similar analysis was done for success in the full course. The results are depicted in table 3.20.

Table 3.20: Student performance in full course according to gender

Gender	Frequency	%	Duration for Full course Success (Semesters)			Total Number of Successful Students	First-time pass rate (Percentage)	Pass Rate as Percentage of Successful students per Gender Group
			3	4	>4			
Female	17	2,9	2	1	0	3	11,8	17,7
Male	567	97,1	32	32	16	70	5,6	12,4
Total	584	100,0	34	33	16	73	5,8	12,5

Female students generally performed better in the full course than their male counterparts. The pass rate for the female students was nearly 50% higher than that of the male students, while their first-time pass rate was twice that of the male students.

3.3.9 Distribution of age

Most of the students were relatively young at entry into the Technikon. The majority of students were aged between 17 and 22, with a mean age of 20 years.

Table 3.21: Age distribution of first semester entrants from 1988 to 1989

Variable	Frequency	Mean	Standard Deviation	Coefficient of Variation	Mode	Range	Median	Q_3-Q_1
Age	584	19.98	3.29	16.47	18	24	19	3

3.3.10 Distribution of individual matric subjects

The marginal distribution of individual matric subjects is given in table 3.22:

Table 3.22: Student performance in individual matric subjects

Variable	Frequency	Mean	Standard Deviation	Coefficient of Variation	Mode	Range	Median	Q_3-Q_1
Afrikaans	516	50,71	9,25	18,24	45	80	45	10
English	548	51,16	9,98	19,50	45	80	45	10
Black	105	57,05	8,76	15,36	55	45	55	10
Maths	548	53,46	12,78	23,90	45	80	55	20
Physical Science	490	49,92	11,40	22,83	45	80	45	10
Biology	461	54,06	11,53	21,33	55	80	55	20
Geography	233	51,13	9,63	18,83	45	44	45	10
History	79	54,31	13,71	25,25	45	80	55	20
Biblical Studies	9	53,11	14,37	27,06	55	38,5	55	10
Accounting	104	51,43	11,63	22,62	45	43,5	55	15
Business Economics	17	51,74	11,31	21,86	45	38,5	55	10
Economics	9	46,61	9,96	21,37	36,5	28,5	45	18,5
Woodwork	18	60,92	12,92	21,21	55	53,5	55	20
Electrician's Work	25	56,72	10,43	18,39	55	38,5	55	10
Fitting and Turning	3	51,67	5,77	11,17	55	10	55	10
Metalwork	12	55	7,39	13,43	55	20	55	10
Technical Drawings	55	54,53	16,58	30,40	45	80	55	20
Agricultural Science	4	50	5,77	11,55	45	10	50	10
Aggregate	551	52,22	7,78	14,90	55	38,5	55	10

The matric subjects offered by the majority of students were the Official Languages, Mathematics, Physical Science, Biology and to a lesser extent Geography, Accountancy and a Black language. The performance of the students in these subjects ranged from approximately 50% for Physical Science to 57% for the Black languages. The variation in marks within individual subjects was relatively low with coefficients of variation ranging from approximately 15,4% for the Black Languages to 23,9% for Mathematics. Very little variation (CV = 15%) around a mean of 52,2% was found for the aggregate marks.

3.3.11 Distribution of Tertiary Study Record:

An analysis of the tertiary study records is given in table 3.23.

Table 3.23 Student performance in individual Technikon subjects

Study Course and Level		Frequency	Mean	Standard Deviation	Coefficient of Variation (CV)	Mode	Range	Median	Q_3-Q_1
Digital Systems	1	284	49,83	21,87	43,90	0	90	54	18,5
	2	130	56,31	12,01	21,50	50	82	56	13
	3	66	55,68	16,71	30,02	65	92	59	14
Electronics	1	300	48,32	22,90	47,38	50	93	52	19,5
	2	167	50,93	17,50	34,35	50	86	53	10
	3	57	52,53	16,47	31,36	53	93	53	15
Electrical	1	370	43,21	22,90	62,73	0	93	51	35
Engineering	2	177	54,02	17,50	33,14	50	90	55	15
	3	40	56,68	16,47	36,08	59	92	59	13,5
Industrial	1	436	53,26	27,11	44,70	0	93	57	18
Technology	2	171	53,02	17,90	35,60	50	84	57	14
	3	25	65,76	20,45	12,67	68	36	68	11
Engineering	1	521	53,26	23,81	53,06	0	98	56	29
Mathematics	2	317	53,08	18,87	33,27	50	99	58	19
	3	184	60,90	8,33	29,71	52	99	61,5	22
Communication									
Language	1	521	50,95	19,66	38,59	0	81	56	11
Communication									
Programming	1	525	55,78	23,27	41,72	0	95	60	19
Telecommun's	2	61	58,56	14,53	24,81	50	83	61	17
Electrical									
Measurements	3	81	54,96	12,81	23,31	50	80	55	11
Radio									
Engineering	3	89	55,98	16,50	29,47	50	88	57	22
Television	3	39	52,49	14,13	26,92	50	77	54	9
Industrial									
Instruments	3	37	57,35	14,85	25,89	56	85	56	16
Engineering	1	125	45,57	21,79	47,81	0	80	51	24
Mechanics	2	49	56,92	13,36	23,47	57	76	58	16
	3	18	69,28	22,92	33,09	89	96	72	28
Electrical	1	278	47,87	27,19	56,80	0	99	53	32
Machines	2	48	53,71	17,84	33,21	50	86	54	13,5
	3	34	51,56	15,39	29,80	50	80	52	21
Strength of	2	178	51,53	17,53	34,01	50	87	54	12
Materials	3	88	57,69	18,29	31,71	50	94	58	19
Mechanical	2	14	70	12,42	17,74	64	45	72	14
Technology	3	1	50	-	-	50	-	50	-
Auto Control	3	40	58,65	21,50	36,66	0	91	63,5	21,5
Engineering									
Science	1	209	51,53	21,56	41,83	0	91	55	17
Machine	1	223	44,91	24,01	53,47	0	91	50	25
Drawings	2	111	52,68	19,13	36,31	50	81	54	16
	3	43	50,28	15,20	30,23	58	70	55	18
Machine Design	2	153	47,54	22,44	47,53	0	97	52	21
	3	61	52,93	17,92	32,53	50	84	58	16
Fluid Mechanics	2	110	46,54	23,42	50,32	50	87	51	18
	3	54	52,93	19,20	36,29	0	86	54,5	18
Mechanics of	2	148	48,55	21,23	43,73	50	91	58	19,5
Machines	3	72	55,46	17,15	30,92	50	92	56	14
Applied									
Thermodynamics	3	499	54,30	17,82	32,82	65	92	58	21
Average		213	55,40	13,17	23,77	57	74	56	16

The average marks attained during the first semester of study ranged from 43,2% in Electrical Engineering with a coefficient of variation of approximately 63%, to 55,8% in Communication (languages) with coefficient of variation of approximately 42%. First semester performance of students was thus average characterised by high levels of variation within individual subjects. During the second academic period student performance showed a slight improvement over the first semester marks. Marks ranged from approximately 47% in Mechanics of Machines (CV = 50%) to 70% (CV = ±18%) in Mechanical Technology. Candidate performance during the final examinations ranged from an average of approximately 50% (CV = 30%) in Machine Design to 69% (CV = 33%) in Industrial Instruments.

When student performance within individual subjects from first to final semester of study was considered, there appeared to be a general tendency of improvement over the period of study. There also seemed to be less fluctuation in student performance as they progressed to more advanced studies. This was evident from a general downward tendency in the coefficient of variation with an increase in academic seniority.

Chapter 4

Chaid Analysis Predictor Selection

4.1 Introduction

The aim of this chapter is to apply the method of Chaid Analysis to the data to identify important predictors of success in student performance in Electrical and Mechanical Engineering at the Peninsula Technikon. Chaid Analysis was selected because of its ability to handle large sets of categorical data. Student performance, at both first and final semester level, was classified into two categories, namely successful and unsuccessful. A student was considered successful at first semester level if he/she obtained a pass-mark of at least 50% in each of the first semester subjects. Success in the full course implied that all requirements for completion of the theoretical component of the full course have been satisfied. Should he/she fail even one of the required subjects, the student was considered unsuccessful. In our attempt to model student performance certain pre-Technikon factors were considered (refer section 2.3).

4.2 Summary

The following observations regarding first semester success were made:

1. The most important predictor of success for those students who wrote the aptitude test was the Swedish Rating scores.

2. The expected pass rate for this group at Swedish rating levels lower than 31 was approximately 30%. The pass rate however improved to approximately 63% at Swedish rating levels above 31.

3. The only other predictor of importance for this group was the student's activity before entering Peninsula Technikon. This was however only the case at Swedish rating levels below 31.

4. Students with tertiary education experience before starting their studies, had a much better chance of success than others. This group had an expected pass rate of 55% against an expected pass rate of 21% for other students.

5. For those students who gained automatic entrance the Swedish rating again emerged as the most important predictor.

6. The expected pass rates for this group improved with improved swedish rating levels in the following way:

   An expected pass rate of 32% at Swedish rating levels below 31;

   An expected pass rate of 57% at Swedish rating levels between 31 and 40;

   An expected pass rate of 86% at Swedish rating levels above 40.

7. For students with Swedish rating scores below 40, the students' activity before registration at Peninsula Technikon was the only other significant predictor.

8. At Swedish rating levels below 30, the expected pass rate for students with previous tertiary education experience was 56%, while only 23% of the other students were expected to be successful. For those in the 31 - 40 Swedish rating group, the expected pass rate was 43% for the group that completed standard 10 during the year before first registering at Peninsula Technikon. The pass rate increased to 77% for the other students. Students that completed matric the year before registration at Peninsula Technikon is thus expected to perform worse than other students.

The results obtained above verified some of the more important preliminary observations made during the exploratory analysis phase. The following observations were made regarding student performance in the full course:

1. Biology emerged as the best predictor of success for those students who wrote the aptitude test.

2. The expected pass rates improved from 4,5% for those who obtained less than 60% in Biology to 19% for those who obtained more than 60%.

3. The Swedish rating emerged as the most important predictor of success for those students who gained automatic entrance.

4. For students in this group the expected pass rates for those with a Swedish rating level below 30 were 4%, and 18% for those with Swedish rating scores above 30.

5. At the lower end of the Swedish rating scale, Biology emerged as the second most important predictor.

6. At these levels student performance improved from approximately 2% to 21% as performance in Biology improved from below to above the 60% level.

The Swedish rating and Biology performance thus emerged as the best predictors of student success.

4.3 Predictor Selection

Chaid Analysis was developed to analyse large sets of data where both the dependent and independent variables are categorical in nature. According to Stoker *et al* (1985: 101) the technique is used mainly to reduce, for a given categorical dependent variable, the dimensions of the multidimensional contingency table that is formed by cross-classifying the categories of the dependent variable with the categories of the independent categorical variables. This reduction can occur in two ways:

1. Determine which categories if any, of the independent variables can be combined.

2. Identify the independent variables that contribute in a statistically significant way towards explaining the variation in the dependent variable.

Once the set of independent variables that have a statistically significant effect on the dependent variable, has been determined, one can then search for a logistic type of regression equation for the dependent variable on this set of predictors. This regression equation can then be used to predict the expected value of the dependent

variable for given values of the predictors. A Chaid Analysis involves the following steps:

a) Stratify each predictor in respect of the dependent variable. This is done by checking and regrouping the categories of a particular predictor into a number of classes, each of which is approximately homogenous in respect of the Y values. Suppose for example, the four categories of a predictor are indicated by the symbols 1, 2, 3 and 4. These categories can for instance be reduced to two classes, namely 1,2 and 3,4. This is done in such a way that interclass differences are significant but not differences amongst the categories within each of the classes. This procedure is followed for each of the predictors.

b) Once this has been done, that predictor which accounts for most of the variation in the Y category values is used in order to divide the data set into a number of subsets.

Each subset is then in turn analyzed in terms of steps (a) and (b) above. This process is continued until no further statistically significant division of the data into subsets is possible. A schematic representation of this partitioning of a data set is known as a dendogram.

4.3.1 Predictor Selection for First Semester Performance

Chaid Analysis was applied to the data to determine the influence of eighteen predictor variables on first semester success. These variables included personal characteristics, matric results and aptitude test results. The variables are detailed below:

Dependent Variable: First semester success.

Predictors: Age, Matric year, Activity previous year, Matric type, Examining body, Home-language, Swedish rating, Mental Alertness, Mechanical Comprehension, Reading Ability, Gottschaldt, Afrikaans, English, Mathematics, Physical Science, Biology and Aggregate.

Two major groups of possible students could be identified at selection. The first group included those students who gained automatic entrance because of their Swedish ratings or some other criteria. The second group comprised those students who because of their low Swedish ratings, were obliged to write the aptitude test. All possible predictors of success available for these two groups were included in the Chaid Analyses. The initial analysis was performed on all the predictors from the group who wrote the aptitude test. This comprised all predictors stated above. This was followed by an analysis of all the possible predictors from the group who gained automatic entrance. The rationale for selecting the two combinations was twofold. Firstly, only selected students wrote the aptitude tests, and secondly a totally different set of predictors than those found for the group that gained automatic entrance might be defining student success for this group.

4.3.1.1 Option 1 - Analysis of matric and aptitude test predictors

First, an analysis was performed on the predictor variables of the group who wrote the aptitude test. This analysis included all 18 variables. The results of the first step in the Chaid process are depicted in table 4.1.

Table 4.1 Predictor variables of first semester success given in order of their predictive power

Predictor	Significance (Ö)	Bonferoni-Significance (Ö)	Number of Groups	Groups
Swedish rating	$5{,}4812 \times 10^{-5}$	0,0001644	2	1,2 3,4
Activity last	0,0038074	0,0266516	2	1,2 3,4
Home Language	0,2486411	1,7404852	2	1 2,3,4
Aggregate	0,4608768	0,9217534	2	1,2 3
Biology	1,3512545	4,0537596	2	1,2,3 4
Physics	1,7564335	5,2692957	2	1,2 3,4
Mental Alertness	2,5121098	7,5363226	2	1,2 3
Examining Body	3,0038404	9,0115175	2	1 2,3
Matric Year	3,7322235	11,1966448	2	1,2 3
Matric Type	4,0225124	60,3375549	2	1,3,5 2,4
Gottschaldt	4,9088593	14,7265425	2	1 2,3
Age	100,00	100,00	1	1,2,3
Reading Ability	100,00	100,00	1	1,2,3
Mech. Comprehension	100,00	100,00	1	1,2,3
Total	100,00	100,00	1	1,2,3
Afrikaans	100,00	100,00	1	1,2,3,4
English	100,00	100,00	1	1,2,3,4
Mathematics	100,00	100,00	1	1,2,3,4

The Swedish rating, with a significance level of $5{,}481 \times 10^{-5}$, was the best predictor of first semester success at the first level of analysis. The first semester performance of students with respect to specific Swedish rating levels are depicted in table 4.2.

Table 4.2: First semester performance according to Swedish rating classes

First Semester Performance	Swedish Rating		Total
	10 - 30	31 - 57	
Fail	71,1%	37,4%	55,9%
Pass	28,9%	62,6%	44,1%
Frequency	121	99	220

The data set was divided into two broad classes, namely those with Swedish ratings below or equal to 30 and those with ratings above 30. Approximately 29% of the 121 students in the 10 - 30 class successfully completed their first semester studies. The corresponding pass rate for the 31 - 57 class was 63,6%. Each of the 10 - 30 and 31 - 57 classes was then separately analysed. First, a summary of the 10 - 30 class analysis is given. This is followed by an analysis of the 31 - 57 class.

4.3.1.2 Summary of analysis of the 10 - 30 class

A summary of the important predictors of first semester performance is depicted in table 4.3. Variables listed in this table are given in order of their importance in explaining first semester performance.

Table 4.3 Predictor variables of the 10 - 30 Swedish rating class

Predictor Variable	Significance	Bonferoni Significance	Number of groups	Groups
Activity last	0,0350233	0,2451627	2	1 243
Home language	0,2633801	1,8436584	2	1 234
Physics	1,0615301	3,1845865	2	123 4

Activity before entering Peninsula Technikon emerged as the best predictor of performance at the second level of analysis. A breakdown of first semester performance in terms of this predictor is depicted in table 4.4.

Table 4.4 First semester performance at the 10 - 30 Swedish rating level according to activity before first registration at Peninsula Technikon

First Semester Performance	Activity before entering Technikon		Total
	1	2,3,4	
Fail	44,8	79,3	71,1
Pass	55,2	20,7	28,9
Frequency	29	92	121

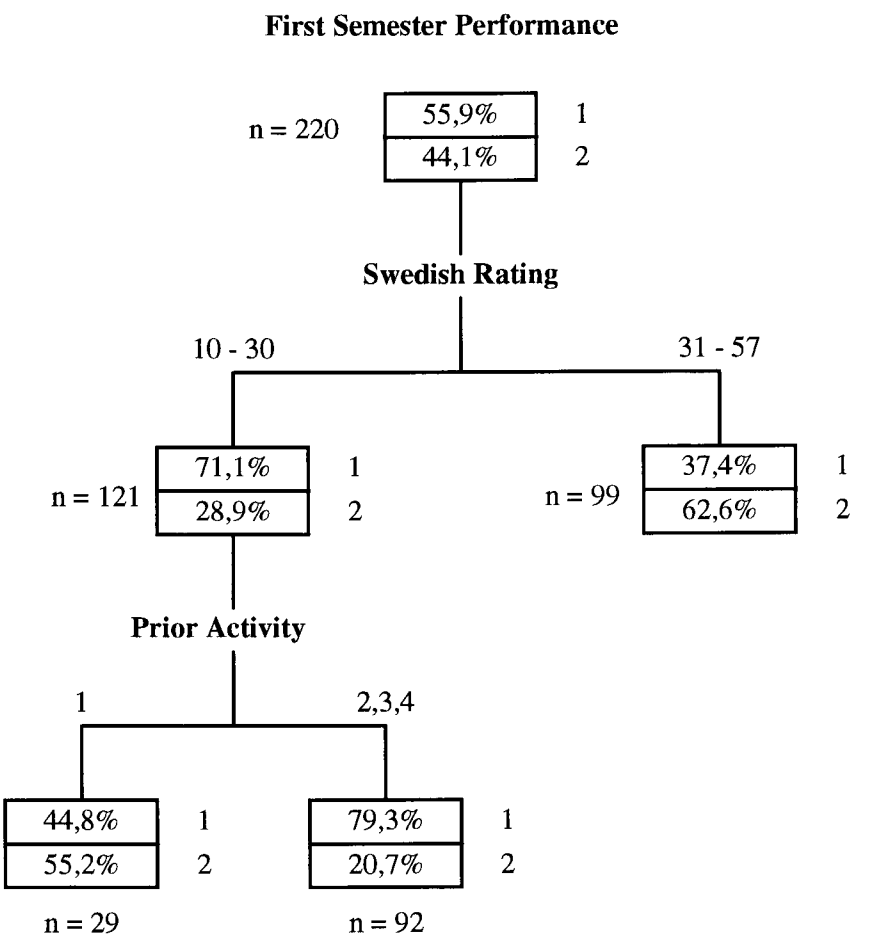
The Chaid analysis divided the data into two classes, namely class 1 and class 2. Class 1 consisted of the University, Technikon and Technical college student categories. The combined class, class 2,3,4, consisted of the student categories Labour force, Standard 10 pupils and the group termed Other. The two groupings are therefore distinguished by whether or not the students have had prior tertiary education experience or not. While 55,2% of the students with prior tertiary education experience successfully completed the first semester studies, only 20,7% of those without previous tertiary education experience were successful. This indicates that students with prior experience performed nearly three times better than

those without. Furthermore, only 28,9% of the 121 students were successful. Since the sample size of class 1 was low, no further analysis was possible. Although the analysis of the 2,3,4 class yielded Physics, examining body and home language as predictors, none of these were statistically significant and no further analysis was performed.

4.3.1.3 Summary of analysis of the 31 - 57 class.

Home language was the most important predictor of success at this level. However, this predictor was not statistically significant at the 5% level and no further analysis was therefore undertaken. Without any significant predictors remaining, further division of the data set was impossible. Combining the results obtained in the foregoing analysis the following dendogram was compiled.

Figure 4.5 Dendogram of First Semester Performance (Option 1)



4.3.1.4 Interpretation of dendogram:

Only 44,1% of the 220 students involved in the analysis successfully completed their first semester studies. Division of this group of students according to the significant predictors of success resulted in the following breakdown.

The model indicated that the predictor, Swedish rating, explained more of the variation in first semester success than any other predictor included in the analysis. At the first level of analysis, the Swedish rating was the best predictor of first semester success. Students with Swedish rating scores of less than 31, had an expected pass rate, for completion of the first semester studies, of 28,9%. The pass rate more than doubled to 62,6% for those students with scores of 31 or greater. This indicates an increase in the expected pass-rate with improved Swedish rating scores.

The two Swedish rating classes, the below 31 and the above 30 classes, are statistically different in respect of the way in which each explained the variation in first semester performance. The second most important predictor of student success was the students' activity before registration at Peninsula Technikon. This predictor was however only significant at the below 31 Swedish rating level. At this Swedish rating level we expect 55,2% of the students with prior tertiary education experience to pass as against only 20,7% of the students without such experience. At Swedish rating scores greater than 30, no statistically significant predictors were identified.

All other predictors were excluded from the dendogram since their contribution in explaining first semester success were not statistically significant.

4.3.1.5 Option 2 - Analysis of matric predictors

The predictors used in the second option were those obtained for the group who gained automatic entrance and hence did not write the aptitude test. At the first level of analysis Swedish rating emerged as the best predictor of success. Three distinct classes of performance levels of this predictor were established, namely those with Swedish rating scores less than 31, the 31 - 40 group and those with scores greater than 40. Student performance at these levels are indicated in table 4.6. It should be borne in mind that this group comprised of students who gained entrance on the strength of their Swedish rating scores or barring this, by virtue of their performance in Mathematics and either Physical- or Applied Science. This would explain the apparent contradiction with respect to students with Swedish scores of 10 - 30 that gained automatic entrance.

Table 4.6 First semester performance according to Swedish rating scores

First semester performance	Swedish Rating			Total
	10 - 30	31 - 40	41 - 57	
Fail	67,8	43,0	14,3	53,5
Pass	32,2	57,0	85,7	46,5
Frequency	180	114	35	329

A vast improvement in student performance with improved Swedish rating scores was observed. The pass rate improved from 32,2% at the below 31 level to 85,7% for those with Swedish rating scores of 41 and above. Each of the three classes was subsequently analysed.

4.3.1.6 Results of the 10 - 30 Swedish rating class.

Analysis of the 10 - 30 Swedish rating class identified activity before registration as the best predictor of success. A breakdown of performance in terms of this predictor is depicted in table 4.7.

Table 4.7 **First semester performance at the 10 - 30 Swedish rating level according to activity before first registration at Peninsula Technikon**

First semester performance	Activity before entering Peninsula Technikon		Total
	1	2,3,4	
Fail	44,0	76,9	67,8
Pass	56,0	23,1	32,8
Frequency	50	130	180

The table indicates that of the 50 students with previous tertiary education experience (class 1) 56% were successful. This was more than twice the pass rate of 23,1% attained by students without such experience, (class 2,3,4). Those class 1 students with Swedish rating between 10 and 30 thus have a much better chance of success than students from the 2,3,4 class. Although certain predictors were identified at the next level of analysis, none were statistically significant. Hence, no further analysis of the data was possible. The 31 - 40 Swedish rating class was then analysed.

4.3.1.7 Results of the 31 - 40 Swedish rating group analysis

Of the possible predictors at the 31 - 40 Swedish rating level, the best predictor of success was again the students' activity immediately before registration at Peninsula Technikon. First semester performance at this level in terms of prior activity is depicted in table 4.8.

Table 4.8 **First semester performance at the 31- 40 Swedish rating level according to activity before first registration at Peninsula Technikon**

First semester performance	Activity before entering Peninsula Technikon		Total
	1,4,2	3	
Fail	23,4%	56,7%	43,0%
Pass	76,6%	43,3%	57,0%
Frequency	47	67	114

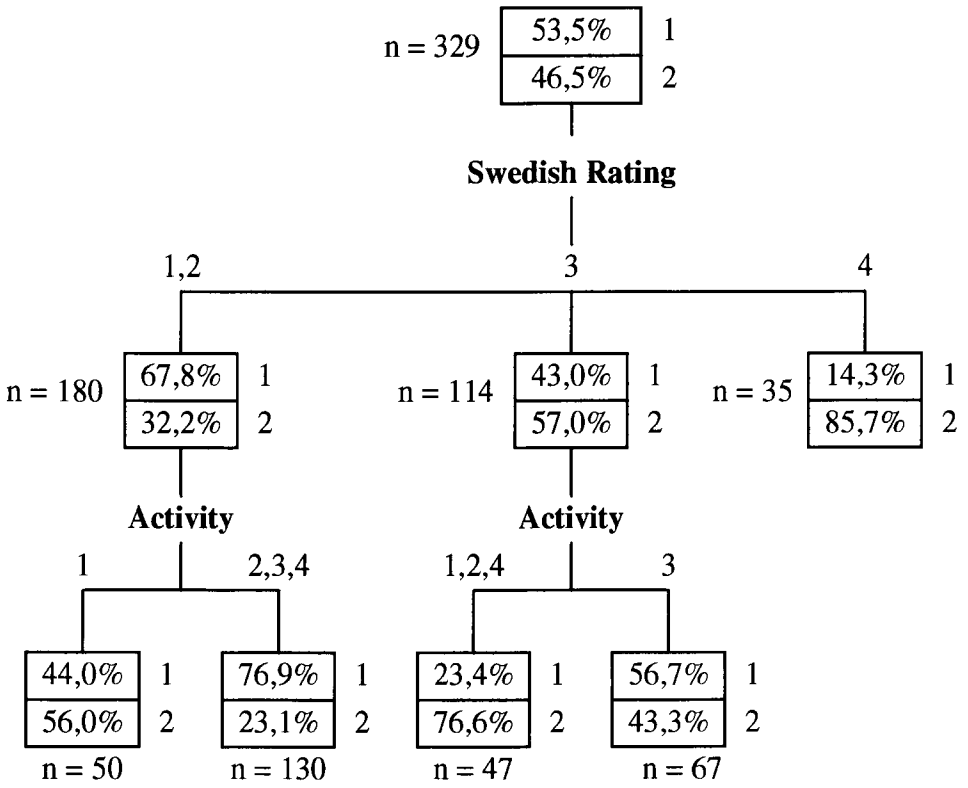
At this level the data was divided into a combined class, class 1,4,2, and class 3. Class 3 was the students who completed standard 10 the previous year, while all other students were grouped together in class 1,4,2. The results indicate that the majority of successful students originated in the 1,4,2 class. In fact, while 76,6% of the students in this class were successful, the corresponding figure for those students directly from school, was 43,3%. Each of these two classes was then further analysed. However, none of the predictors identified in either of the analyses were statistically significant. Hence no further analysis of the data in either of the two classes was undertaken.

Whereas the data in the 10 - 30 class was divided into two classes of students based on previous tertiary education experience, the data in the 31 - 40 class was divided on the basis of whether students were matriculants or not in the year before registration at Peninsula Technikon. This tendency indicates that the predictor, activity before entering Peninsula Technikon, impacts differently on student performance at different Swedish rating levels. Analysis of the third Swedish rating class yielded the following results.

4.3.1.8 Results of the 41 - 57 Swedish rating class analysis

Although the examining body was identified as the best predictor of success at this level, it was not statistically significant and was therefore not included as a predictor. No further analysis of this class was possible. A combination of the results of the analyses above yielded the following dendogram.

Figure 4.9 Dendogram of First Semester Performance (Option 2)



4.3.1.9 Interpretation of dendogram:

Fewer than 50% of the 329 students involved in the analysis successfully completed their first semester studies. The model indicates that the predictor, Swedish rating, again explained more of the variation in first semester success than any other predictor included in the analysis. At the first level of analysis, Swedish rating therefore emerged as the best predictor of success. Three distinct performance classes within this variable were identified.

The dendogram suggests that a considerable increase in pass-rate with improved Swedish rating scores could be expected at this level. According to the diagram an expected 32,2% of the students with Swedish rating scores less than 31 would be

successful, while the expected pass rate increased to 57,0% for those students with scores between 31 and 40 and finally to 85,7% for those with scores of 41 or higher. The second most important predictor was the candidates' activity before registration at Peninsula Technikon. This predictor was however only significant at Swedish rating levels 10 - 30 and 31 - 40. At these Swedish rating levels two distinctly different groupings of prior activity classes were identified as being significant in explaining success. These were classes 1 and 2,3,4 at the 10 - 30 level and classes 1,2,4 and 3 at the 31 - 44 level.

At the 10 - 30 Swedish rating level we expect 56% of the class 1 students to pass while only 23,1% of the class 2,3,4 students are expected to be successful. We furthermore expect 76,6% of the class 1,2,4 candidates and 43,3% of the class 3 candidates at the 31 - 40 level to be successful. At Swedish rating levels above 40, no statistically significant predictor could be identified.

No other predictor was included in the dendogram since their contributions in explaining first semester success were not statistically significant.

4.3.2 Predictor Selection for Final Semester Performance

The same combinations of predictors used during the selection process for first semester success were again used. The initial Chaid analysis was performed on the 18 predictors obtained from the group that wrote the aptitude test and yielded the following results.

4.3.2.1 Option 1 - Analysis of matric and aptitude test predictors

Biology emerged as the best of the three possible predictors identified at the first level of analysis. Student performance in the final semester in terms of this predictor is given in table 4.10.

Table 4.10 Final semester performance according to performance in Biology

Final Semester Performance	Biology		Total
	0 - 60	> 60	
Fail	95,5%	81,0%	91,4%
Pass	4,5%	19,0%	8,6%
Frequency	157	63	220

Low pass rates in both classes were the outstanding characteristic of the results. The pass rate in the below 60% class was in fact only 4,5%. The students with Biology marks greater than 60% were nearly four times as successful as those in the below 60% class. Analysis of each of the two Biology classes was subsequently done. Analysis of the 0 - 60 class yielded the following results.

4.3.2.2 Results of the 0 - 60 class analysis

At the second level of analysis the type of Matriculation qualification emerged as the best predictor of final semester performance. Final semester performance according to the type of matriculation qualification is given in table 4.11.

Table 4.11 Final semester performance according to type of matric qualification

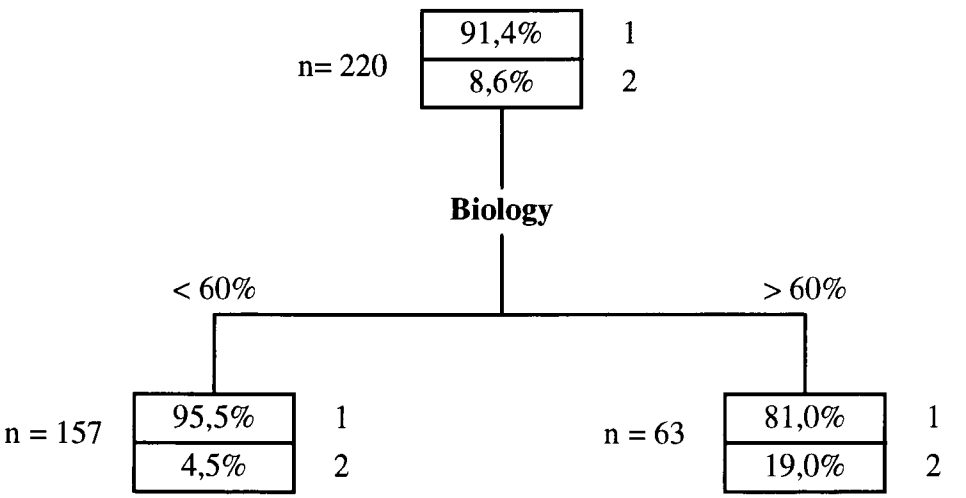
Final Semester Performance	Matric type		Total
	1,2,4,5	3	
Fail	96,8%	0%	95,5%
Pass	3,2%	100,0%	4,5%
Frequency	155	2	157

Student performance was defined in terms of two classes, namely the combined 1,2,4,5 class and class 3. Class 3 consisted of the students with a Technical college qualification, while the rest of the students were grouped into the combined class. Since no significant conclusions could be drawn with only two observations in class 3, a comparative analysis of the two groups would be a futile exercise. It is however worthwhile to note that only 3,2% of the 155 students from the combined group was successful. Analysis of the above 60 class yielded the results below.

4.3.2.3 Results of the 60+ class analysis

Since no statistically significant predictor was found at this level, the selection process was concluded. The following dendogram was subsequently compiled from the results of the foregoing analyses.

Figure 4.12 Dendogram of Final Semester Performance (Option 1)



4.3.2.4 Interpretation of the dendogram

Of the 220 students involved in the analysis, only 8,6% successfully completed the full course. The best predictor of success was student performance in Biology. Two performance classes of the predictor were identified, namely those with performance levels below 60%, and those above. At a performance level below 60%, only 4,5% of prospective students are expected to be successful. The pass rate was expected to increase fourfold at Biology performance levels higher than 60%.

4.3.2.5 Option 2 - Analysis of matric predictors

The second Chaid analysis was performed on data obtained from the group that gained automatic entrance. The aptitude test results were thus omitted from the list of possible predictors used in the selection process. At the first level of analysis the Swedish rating emerged as the best predictor of success. Final semester performance according to the identified Swedish rating classes is given below.

Table 4.13 Final semester performance according to Swedish rating scores

Final Semester Performance	Swedish Rating		Total
	11 - 30	31 - 57	
Fail	96,1%	82,5%	89,8%
Pass	3,9%	17,5%	10,2%
Frequency	180	154	334

The data was stratified into two classes namely, the 11 - 30 and the 31 - 57 Swedish rating classes. Although the 31 - 57 class had a pass rate of more than four times that of the students in the 11 - 30 class, its pass rate was nevertheless very low at 17,5%. These Swedish rating classes were subsequently analysed, the results of which follows.

4.3.2.6 Results of the 11 - 30 Swedish rating class analysis

The best predictor of success was Biology that was stratified into a below 60% and an above 60% class. Student performance according to these two classes is given in table 4.14.

Table 4.14 Final semester performance at the 11 - 30 Swedish rating level according to performance in Biology

Final Semester Performance	Biology		Total
	< 60%	> 60%	
Fail	98,1%	78,9%	96,1%
Pass	1,9%	21,1%	3,9%
Frequency	161	19	180

The pass rate of 21,1% at the above 60% level was nearly 11 times that of the below 60% class. However, the significance of this observation is compromised by the low sample size of the above 60% class. The pass rate of 1,9% for the below 60% class was very low. Further analysis was only performed on the below 60% class since the sample size of the above 60% class was too low. The best predictor at this level was the type of Matriculation qualification. Results of this analysis are given below.

Table 4.15 Final semester performance at the below 60% performance level in Biology according to matriculation type

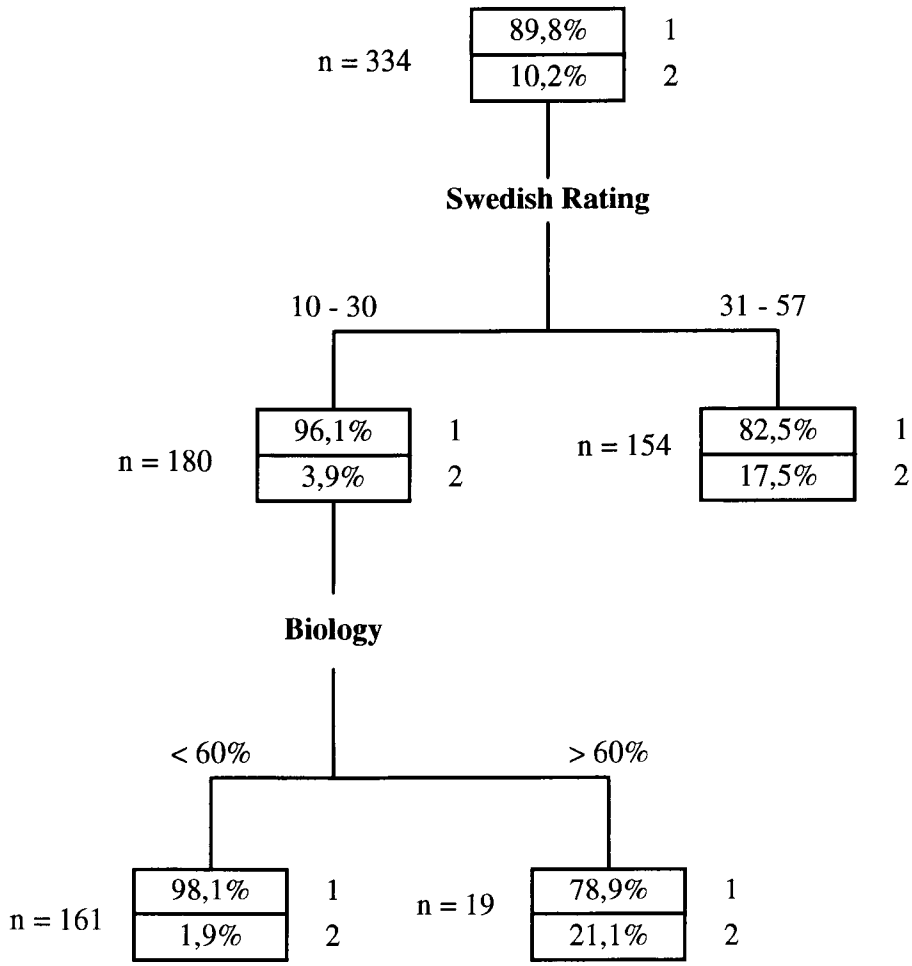
Final Semester Performance	Matriculation Type		Total
	1,4,5,2	3	
Fail	98,7%	50,0%	98,1%
Pass	1,3%	50,0%	1,9%
Frequency	159	2	161

The pass rate of 1,3% at the 1,4,5,2 class was very low. Since there were only two students in class 3, no significant deductions regarding this group or any comparative analysis with the 1,4,2,5 class were possible. Hence all analyses were concluded. Analysis of the 31 - 57 Swedish rating class yielded the results below.

4.3.2.7 Results of the 31 - 57 Swedish rating class analysis

The best predictor of success at this level of analysis was age, but because this predictor was not statistically significant, it was excluded as a predictor of success. No further analysis was thus undertaken. The following dendogram was compiled from the results obtained.

Figure 4.16 Dendogram of Final Semester Performance (Option 2)



4.3.2.8 Interpretation of Dendogram

At the first level, Swedish rating emerged as the best predictor. There was a fourfold increase in the expected pass rate of students with improved Swedish rating levels. Thus we expect approximately 4% of the 10 - 30 class students to pass as opposed to approximately 18% of the 31 - 57 class students.

At the 10 - 30 Swedish rating level, Biology emerged as the best predictor of student performance. Approximately 2% of students with pass mark below 60% are expected to be successful. The expected pass rate improved to 21% for students

with pass marks greater than 60%. We therefore expect a considerable increase in pass rate with improved Biology performance.

At the 31 - 57 Swedish rating level, the predictors identified were not statistically significant and this class was thus not further analysed.

In conclusion, the two predictors that were significant in predicting student performance in the full course were the students' scores on the Swedish rating and their performance in Biology.

Chapter 5

Discriminant Analysis

5.1 Introduction

In this chapter we use the method of Discriminant Analysis to assess the predictive nature of the variables in the data set. In particular this technique assists in establishing the important predictors of successful or unsuccessful studies from sets of variables defined on these two groups. Successful and unsuccessful studies were defined as in section 4.1.

5.2 Summary

The following observations regarding first semester performance were made:

1. For the students who wrote the aptitude test, a 6-variable model, consisting of the Swedish rating, Technical reading, Mechanical comprehension, Afrikaans, Mathematics and English, was defined as the model most efficient in discriminating between successful and unsuccessful students.

2. Of these predictors, the Swedish rating was the best.

3. Should this model be used in the student selection process, 68% of the selected students would be expected to be successful, while approximately 31% of the rejected students would be expected to pass.

4. For the students who gained automatic entrance, a 3-variable model comprising the Swedish rating, Mathematics and Afrikaans emerged as the model best able to discriminate between success or failure.

5. The Swedish rating was again the best predictor.

6. If this model were used in the selection process, the expected pass rate of those selected would be 60%, while approximately 28% of the rejected students would have passed if they were selected.

7. Both these models are expected to yield pass rates well above that of the current selection model.

The following observations regarding performance in the full course were made:

1. For the students who wrote the aptitude test, a 3-variable model comprising Biology, Swedish rating and English, was developed as the model best able to discriminate between potentially successful and unsuccessful students.

2. Biology emerged as the best predictor, followed by the Swedish rating and English.

3. Should this model be used in the student selection process, none of the students selected would be expected to pass. However, the significance of this observation is minimal since it is based on the results of a single student. Approximately 12% of the rejected group would be expected to pass if they were selected.

4. For the students who gained automatic entrance, a 2-variable model comprising the Swedish rating and Physical Science was developed.

5. If we use this as the student selection model, we expect 4% of the selected students and 1% of the rejected students to pass.

6. Both models yielded expected pass rates well below those attained through the current model.

5.3 Discriminant Analysis

According to Mardia *et al* (1979:300), the primary aim of the discrimination process is the allocation of observations to one of a number of distinct groups. In particular, given measurements of these observations on a number of variables, this process focus on how well it is able to do the allocation.

This method allows us to search for a rule that will allow us to adequately discriminate between the groups. With data on m successful and n unsuccessful students, for example, it is useful to find a way of adequately discriminating between the successful and unsuccessful students.

In order to decide which group an arbitrary observation belongs to, one needs to ask how close that observation is from any of the two groups. One way of estimating this "closeness" is by means of the Mahalanobis distances. The Mahalanobis distance is used to measure the distance of a single multivariate observation from the centre of a population from which the observation originates. Once this is done, the observation can be allocated to the group that it is closest to. This may or may not be the group that the individual actually came from.

If an observation is allocated to the group from which it originates, the allocation is said to be correct, while an observation allocated to a group other than that from which it originates, is considered as a wrong allocation. The percentage of correct allocations is clearly an indication of how well groups can be separated using the available variables. This procedure is defined as follows:

Let $\bar{\mathbf{x}}_i' = (\bar{x}_{1i}, \bar{x}_{2i}, \ldots, \bar{x}_{pi})$ denote the vector of mean values for the sample from the i-th group and let C_i denote the covariance matrix for the same sample, and C the pooled sample covariance matrix. Then the Mahalanobis distance from an observation $\bar{x}' = (x_1, x_2, \ldots, x_p)'$ to the centre of group i is estimated as:

$$D_i^2 = (\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)' C^{-1} (\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)$$
$$= \sum_{r=1}^{p} \sum_{s=1}^{p} (\bar{x}_r - \bar{x}_{ri}) c^{rs} (\bar{x}_s - \bar{x}_{si})$$

where c^{rs} is the element in the r-th row and the s-th column of C^{-1}.
The observation is then allocated to the group for which D_i^2 has the smallest value.
It is sometimes useful to be able to determine functions of the variables $X_1, X_2, \ldots, X_p$ that separate the m groups in one or other manner as well as possible.
The implicit approach is to set up the discriminant function, say F, by taking a linear combination of the X variables, e.g.:

$$F = a_1 X_1 + a_2 X_2 + \ldots + a_p X_p$$

where a_i are chosen such that the F - ratio between groups is maximized, i.e. the ratio of the between group variance to the total variance should be as large as possible. While the possibility of misclassifying observations exists, the procedure should result in as few as possible misclassifications. In order to facilitate this, prior probabilities of group membership should be considered.

Given two groups, it might be known that most observations fall within a specific group while fewer observations fall within the second group, i.e. the probability of membership is inherently different for the two groups. It thus makes sense to bias the process of allocation of observations in favour of the larger group. Prior probability of group membership was taken into account in this analysis.

Allocation of observations would naturally be biased in favour of the group from which they derive, since these observations were used to calculate the group mean. Thus observations tend to be "closest" to the centre of the group to which they contributed. To overcome the bias a "jackknife classification" of observations is carried out. This involves calculating a group mean using all observations in the group except the observation under consideration. The observation is then allocated to that group to which centre it is the closest. In this way bias in the allocation of observations can be avoided. This process is repeated for all observations in the sample.

For purposes of developing a statistically significant discriminant function to separate the groups, stepwise discriminant analysis was used. The suitability of using such a procedure is summed up by Hair (1992: p99), who defines stepwise discriminant analysis as a computational method used to derive a discriminant function to separate two or more groups.

5.3.1 Stepwise Discriminant Analysis

This procedure was decided upon since the primary aim of the analysis was to determine which matriculation or aptitude test variables are most efficient in discriminating between successful or unsuccessful candidates. The procedure involved entering independent variables into the discriminant function, one at a time, based on their discriminating power.

This procedure is designed to develop the best one-variable model, followed by the best two-variable model and so forth, until no further variable meets the requirement of maximizing the Mahalanobis distance, D^2, between groups. The process begins with all variables excluded from the model. Variables enter or leave the model based on one of the following criteria:

1. Significance level of an F-test, where the variables already entered act as covariates and the variable under consideration is the dependent variable.

2. The squared partial correlation for predicting the variable under consideration, while controlling for the variables already in the model.

The variable best able to discriminate between the two groups, is entered into the model first. This variable is then paired with each of the remaining variables and a second variable selected. Selection of the second variable depends on which combination of the variable already in the model and one of the remaining variables contributes most to improving the discriminant function's power. All further variable selections are done in the same way.

As more variables are entered into the model previously selected variables may be removed, if the information they contain regarding within-group differences are available in some combination of the other included variables. The selection process stops when inclusion of additional variables does not significantly improve the discriminant function's ability to discriminate between groups.

5.3.2 Selection of predictors of first semester performance

The list of variables from which the predictors were selected comprised both matriculation results as well as aptitude test results.

Only two combinations of possible predictors, based on actual information available at the time of student registration, were considered. The first combination comprised matric results only, while the second consisted of matric results plus aptitude test results. Stepwise discriminant analysis using the forward selection procedure was used to determine the best subset of predictors.

Predictor variables used: Swedish rating, matric aggregate, Mathematics, Physical Science, English, Afrikaans, Biology, Mechanical comprehension, Mental alertness, Technical reading, Gottschaldt figure test and aptitude aggregate.

The dependent variable, first semester performance, was classified either as successful or unsuccessful. The classification was based on the definitions stated in section 4.1.

5.3.2.1 Procedure

The two combinations of variables were individually analysed with the aid of the SAS PROC DISCRIM procedure. The first combination comprised matric as well as aptitude test results. Two hundred and twenty-two observations on twelve variables (see table 5.1) were analysed. Membership of the 222 observations to the two groups of the dependent variable was as follows:

Successful candidates : 98

Unsuccessful candidates : 124

The criterion for variable entry (SLE = significance level to enter), as well as that for the removal of variables (SLS = significance level to stay) was set at 0,15. The selection process was initiated with all possible predictors excluded from the model. All the necessary statistics were then computed and the selection of that single predictor that was most effective in discriminating between the two groups, was allowed to enter the model. Since predictor selection was based on the F-values, the variable with the largest F-value was entered into the 1-variable model first. At this level, the variable entered into the model was the Swedish rating with an F-value of 34,79. The statistics for the removal of variables were subsequently computed to ascertain whether the included variable was significant. Swedish rating was found to be highly significant and thus remained within the model. Since Afrikaans, English and the aptitude aggregate had F-values less than 1, they should have been removed from the predictor list. However, once the effect of the Swedish rating was removed, the F-values of these variables improved to levels significantly above one.

At the second level of analysis the best 2-variable model was selected by pairing the Swedish rating with all the variables outside the model, one at a time. At this level Technical reading with an F-value of 4,709 was entered into the model. The significance of this variable to remain within the model was then verified. The model selection process was continued until no further variables could be included or removed from the model.

5.3.2.2 Results

A summary of the first two steps in the model selection process is reflected in table 5.1.

Table 5.1 Summary of statistics of the first two steps in the stepwise model selection process, using the forward selection procedure.

Step	Variable	Partial R^2	F value	Significance level
1	Afrikaans	0,0012	0,261	0,6099
	English	0,0024	0,530	0,4672
	Mathematics	0,0046	1,025	0,3125
	Physics	0,0241	5,425	0,0208
	Biology	0,0381	8,722	0,0035
	Aggregate	0,0433	9,969	0,0018
	Mental Alertness	0,0065	1,435	0,2322
	Mechanical comprehension	0,0227	5,118	0,0247
	Aptitude aggregate	0,0014	0,316	0,5743
	Technical reading	0,0125	2,794	0,0960
	Gottschaldt	0,0092	2,045	0,1541
	Swedish rating	0,1365	34,786	0,0001
Best 1-variable model: Swedish rating **Average squared Canonical correlation = 0,1365**				
Variable removal statistics				
	Variable	R^2	F value	Significance level
	Swedish rating	0,1365	34,786	0,0001
Variable Swedish rating cannot be removed from model				
2	Afrikaans	0,0200	4,459	0,0358
	English	0,0114	2,523	0,1136
	Mathematics	0,0207	4,628	0,0326
	Physics	0,0017	0,376	0,5406
	Biology	0,0003	0,073	0,7880
	Aggregate	0,0077	1,698	0,1939
	Mental Alertness	0,0001	0,029	0,8653
	Mechanical comprehension	0,0065	1,432	0,2327
	Aptitude aggregate	0,0019	0,426	0,5147
	Technical reading	0,0210	4,709	0,0311
	Gottschaldt	0,0004	0,086	0,7694
Variable Technical reading will be entered into model **Best 2-Variable model: Swedish rating and Technical reading** **Average squared Canonical correlation = 0,1547**				
Variable removal statistics				
	Variable	Partial R^2	F value	Significance level
	Technical reading	0,0210	4,709	0,0311
	Swedish rating	0,1440	36,832	0,0001
None of the variables can be removed				

The average squared canonical correlation values are given at each level of analysis. This statistic is interpreted as the percentage of the variance in the dependent variable explained by the particular model. So, for example, 13,65% of the variance in student performance can be explained by the 1-variable model, 15,47% by the 2-variable model, and so forth. A summary of the results of the full stepwise procedure is given in table 5.2.

Table 5.2: Summary of the two-group stepwise discriminant analysis results.

Step	Variable		Partial R^2	F value	p - valuel	Average squared canonical correlation	p - value	Wilks Lambda	p - value
	Entered	Removed							
1	Swedish rating	-	0,1365	34,786	0,0001	0,1365	0,0001	0,8635	0,0001
2	Technical Reading	-	0,0210	4,709	0,0311	0,1547	0,0001	0,8453	0,0001
3	Mechanical Comprehension	-	0,0207	4,613	0,0328	0,1722	0,0001	0,8278	0,0001
4	Afrikaans	-	0,0261	5,815	0,0167	0,1938	0,0001	0,8062	0,0001
5	Mathematics	-	0,0343	7,665	0,0061	0,2215	0,0001	0,7785	0,0001
6	English	-	0,0104	2,265	0,1338	0,2296	0,0001	0,7704	0,0001

Table 5.2 indicates that six variables, i.e. Swedish rating, Mechanical comprehension, Technical reading, Afrikaans, Mathematics and English entered the model and are all significant predictors (discriminators) based on their Wilks' lambda values. Inasmuch as Wilks' lambda tests for equality of the two group means, the results would indicate that these groups are significantly different in terms of their performance in the indicated predictors. Performance in these variables could therefore play a significant role in the allocation of prospective candidates to one of these groups.

The contribution of each variable in explaining the variance in student performance is reflected by the partial R^2 values. The shared variance in student performance explained by the cumulative effect of included predictors at each step, is described by the average squared canonical correlation. A similar analysis was performed with Matric subjects only as possible predictors of first semester performance. All test parameters were set as for the previous analysis. A summary of the 3-variable model resulting from the analysis is given in table 5.3. Only three of the seven possible predictors were included in the final model that was subsequently used to develop a discriminant function for the purpose of allocating candidates into one of the two performance groups. These were Swedish rating, Mathematics and Afrikaans.

Table 5.3 Summary of Stepwise discriminant analysis results using matric results only

Step	Variable		Partial R^2	F value	p - valuel	Average squared canonical correlation	P - value	Wilks Lambda	p - value
	Entered	Removed							
1	Swedish rating	-	0,1489	57,752	0,0001	0,8511	0,0001	0,1489	0,0001
2	Mathematics	-	0,0154	5,142	0,0240	0,8380	0,0001	0,1620	0,0001
3	Afrikaans	-	0,0239	8,048	0,0048	0,8179	0,0001	0,1821	0,0001

The Wilks' lambda values indicate that all three predictors, Swedish rating, Mathematics and Afrikaans are significant in terms of differentiating between the two first semester performance groups. The respective contributions of these variables in explaining the variance in student performance is indicated by the partial R^2, while the variance explained by the combined effect of the included predictors at each level is indicated by the average squared canonical correlation values. Although this subset of predictors could explain only approximately 18% of the variance in student performance at first semester level, as against the nearly 23% of the previous subset, it is important in at least two ways:

1. Results on fewer variables are needed on which decisions are based.
2. Aptitude results are normally available only for a limited percentage of candidates.

5.3.3 Identification of predictors of final semester performance.

The next step was to ascertain the importance of the same sets of possible predictors on student performance at the final semester level, and whether those subsets of predictors identified as statistically significant at the first semester level were also important at final semester level.

5.3.3.1 Procedure

The same two subsets of predictors used in the analysis of the first semester success were again selected. All first semester test parameters, e.g. SLE and SLS levels, prior probabilities, etc were retained, while the dependent variable was replaced by the students' performance at final semester level. Performance was again divided into the two achievement groups, successful and unsuccessful. The PROC STEPWISE DISCRIM procedure was used to perform the necessary statistical procedures. In option 1 we considered all available data from the group of students who wrote the aptitude test. Option 2 dealt with those students who gained automatic entrance to the course.

5.3.3.2 Results

Results for the stepwise analyses performed are given in table 5.4.

Table 5.4 Summary of the two-group stepwise analysis results for final semester performance.

Option	Variables Used	Number of observations per class			Variables Retained as statistically significant	% Explained variation (ASCC)
		(1) Unsuccessful	(2) Successful	Total		
1	Afrikaans, English, Mathematics, Biology, Physics, Aggregate, Swedish rating, Men, Mech, Read, Gots, Aptitude aggregate	203	19	222	Biology, Swedish rating, English	7,39
2	Afrikaans, English, Mathematics, Biology, Physics, Aggregate and Swedish rating	303	34	337	Swedish rating, Physical science	8,58

Note: *All statistically significant predictors are given in order of their importance in discriminating between the successful and unsuccessful groups.*

When considering all matric and aptitude test variables as possible predictors, Biology, Swedish rating and English were identified as statistically significant. Inclusion of predictors into the final 3-variable model were in the order given in table 5.4. While the variance in final performance explained by the 3-variable model is 7,39%, the corresponding value for the 2-variable model consisting of Swedish rating and Physical science was 8,58%. Both models therefore gave rise to low R^2 values.

5.3.4 Discriminant function development

The next step in the process was developing linear discriminant functions necessary for the allocation of candidates into either the successful or unsuccessful groups, using the subsets of predictors identified during the previous stage. Two separate linear functions, each defining the contribution of the identified predictors in developing an appropriate allocation criterion for the relevant group, were developed. This procedure was excecuted for each combination of significant predictors obtained from the stepwise discriminant procedure, both at first as well as final semester level.

The appropriateness of these functions was then determined on the basis of misclassification results obtained for the respective groups.

5.3.4.1 First Semester discriminant function development procedure

The PROC DISCRIM procedure was used to determine the ability of the two models obtained during the PROC STEPWISE procedure to discriminate between success or failure at first semester level.

5.3.4.1.1 The 6-variable model

The 6-variable model obtained for the first semester performance was the first to be analysed. Information relevant to the process is given below:

Dependent variable : First semester performance

Independent variables : Swedish rating, Technical reading, Mechanical comprehension, Afrikaans, Mathematics and English

Performance group	Frequency	Prior probability	Generalized squared distances
Successful	140	0,4651	1,5309
Unsuccessful	161	0,5349	1,2514

A linear discriminant function for each of the performance groups was subsequently determined. Using the discriminant function, a discriminant score for each observation in the sample was calculated. Observations were then allocated to a particular group on the basis of the calculated scores. The percentage of cases classified correctly according to the obtained discriminant functions was then calculated.

5.3.4.1.2 Results

A summary of the classification results of the discriminant analysis procedure is given in table 5.5.

Table 5.5 Summary of discriminant analysis results for the 6-variable model

	DISCRIMINANT FUNCTION			
	Successful group		Unsuccessful group	
	Variables	Discriminant loading	Variables	Discriminant loading
	Swedcor	-0,3711	Swedcor	-0,5271
	Read	0,6984	Read	0,8448
	Mech	1,9332	Mech	1,6539
	Afr	0,4459	Afr	0,4750
	Maths	0,4420	Maths	0,4600
	Eng	0,4870	Eng	0,5100
Percentage cases correctly classified	61,43 (86)		75,16 (121)	
Total error Rate†	38,57		24,84	

Allocation of students in accordance with the attained discriminant functions leads to the results depicted in the table below.

Table 5.6 First semester performance according to the 6-variable model

		Predicted Status		
		Success	**Failure**	**Total**
Actual Status	**Success**	**86**	**54**	**140**
	Failure	**40**	**121**	**161**
	Total	**126**	**175**	**301**

The table indicates that 86 students from the successful group and 40 students from the unsuccessful group would have been selected onto the course on the basis of the 6-variable model. Of the 126 students selected, the percentage of potentially successful students in the group is 68,3% (86/126), while the potentially successful students in the rejected group is 30,9% (54/175).

5.3.4.1.3 The 3-variable model

The same procedure was followed when analysing first semester performance on the basis of the 3-variable model. A summary of the results of the discriminant procedure is given in table 5.7.

Table 5.7 Summary of discriminant procedure results of student performance at first semester level on the basis of the 3-variable model.

	DISCRIMINANT FUNCTION			
	Successful group		Unsuccessful group	
	Variables	Discriminant loading	Variables	Discriminant loading
	Swedcor Maths Afr	-0,0790 0,3667 0,6256	Swedcor Maths Afr	-0,1910 0,3742 0,6393
Percentage cases correctly classified	59,91 (130)		71,97 (172)	
Total error Rate†	40,09		28,03	

The percentage of candidates correctly classified into the groups on the basis of the 3-variable model is only marginally worse than the corresponding figures attained when using the 6-variable model. Had this model been adopted to select students onto the programs, the following results would have been attained for the given data.

Table 5.8 First semester performance according to the 3-variable model

		Predicted Status		
		Success	Failure	Total
Actual Status	Success	130	67	197
	Failure	87	172	259
	Total	217	239	456

According to table 5.8, 130 students from the successful group and 87 students from the unsuccessful group would have been selected onto the course on the basis of the 3-variable model. Of the 217 students selected, the percentage successful students in the group is 59,9% (130/217), while the successful students in the rejected group is 28,0% (67/239).

5.3.4.2 Interpretation of the discriminant analyses

The results indicate that, should the 6-variable model be applied in the selection process, we would expect:

Approximately 68% of the students admitted onto the course and approximately 31% of the rejected students would be successful. Stated differently, we would have made a correct decision in the selection of successful candidates 68% of the time. While the success rate in identifying unsuccessful candidates is predicted to be 69%, the 31% of successful students who would be incorrectly classified as potential failures is cause for concern.

The 6-variable model is thus not very effective in identifying potentially successful and unsuccessful students. Should the 3-variable model be used as selection tool, we would expect the following results:

Approximately 60% of the accepted students and about 28% of the rejected students, would be successful. Stated differently, we would have made a correct decision in the selection of successful candidates 60% of the time. While the success rate in identifying unsuccessful candidates was 72%, the 28% of successful students who would be incorrectly classified as potential failures is still high.

As was the case with the 6-variable model, the 3-variable model is also not very effective in identifying potentially successful and unsuccessful students. Although the 6-variable model yielded better results in the identification of potentially successfully students, the alternative 3-variable model performed slightly better in the identification of potentially unsuccessful students.
The advantage enjoyed by the 6-variable model over the alternative 3-variable model, when identifying potentially successful students, has to be weighed up against the additional time, effort and resources required to obtain the relevant data.

Thus none of the models provide satisfactory discriminant functions for the correct allocation of candidates to any of the two groups. Both models are however expected to yield pass levels higher than those attained through the current model (compare the expected pass rates of 68% for the 6-variable model and 60% for the 3-variable model against the approximately 49% for the current model).

5.3.4.3 Final Semester discriminant function development procedure

5.3.4.3.1 The 3-Variable model

The PROC DISCRIM procedure was again used to determine the ability of the two models obtained during the PROC STEPWISE procedure to discriminate between success or failure in the final examinations. The variables included in the analysis were:

Dependent variable : Final achievement (Fin)

Predictor variables : Biology, Swedish rating (Swedcor) and English (Eng)

Performance in the final semester was divided into two groups; successful and unsuccessful. Membership of these groups plus other relevant basic statistics obtained were:

Performance group	Frequency	Prior probability	Generalized squared distances
Successful	52	0,1235	4,1828
Unsuccessful	369	0,8765	0,2637

Significance levels for variable entry and removal were fixed at the same levels as for first semester achievement.

A linear discriminant function for each of the performance groups was then determined, followed by the calculation of discriminant scores for each observation in the analysis. Observations were allocated to a particular group on the basis of the calculated scores. The percentage of cases correctly classified, as well as the total error rates for both successful and unsuccessful groups were then determined.

5.3.4.3.2 Results

A summary of the classification results of the discriminant analysis for the 3-variable model is given in table 5.9.

Table 5.9 Summary of discriminant analysis results for the 3-variable model.

DISCRIMINANT FUNCTION				
	Successful group		Unsuccessful group	
	Variables	Discriminant loading	Variables	Discriminant loading
	Biology Swedcor English	0,3697 0,0402 0,4551	Biology Swedcor English	0,3438 -0,0311 0,4669
Percentage cases correctly classified	0,00 (0)		99,73 (368)	
Total error Rate†	100,0		0,27	

The classification function for the unsuccessful group has a very high level of success in correctly classifying unsuccessful candidates (99,7% or 368 out of 369 candidates), while the success rate attained by the successful group's classificatiion function amounted to 0,0% (0 out of 52 candidates). While performance levels in the Swedish rating and Physical science could therefore adequately predict potentially unsuccessful candidates, it is totally inadequate in predicting potential success.

Allocation of students in accordance with the attained discriminant functions would lead to the results depicted in the table below.

Table 5.10 Final semester performance according to the 3-variable model

		Predicted Status		
		Success	**Failure**	**Total**
Actual Status	**Success**	**0**	**52**	**52**
	Failure	**1**	**368**	**369**
	Total	**1**	**420**	**421**

The table indicates that none of the students from the successful group and one of the students from the unsuccessful group would have been selected onto the course on the basis of the 3-variable model. Of the single student selected, the percentage of potentially successful students in the group is 0% (0/1), while the potentially successful students in the rejected group is 12,4% (52/420).

The procedure was then repeated for the identified 2-variable model.

5.3.4.3.3 The 2-Variable model

Performance in the final semester was again divided into the two performance groups: successful and unsuccessful. Membership of these groups plus other relevant basic statistics obtained were:

Performance group	Frequency	Prior probability	Generalized squared distances
Successful	54	0,1236	4,1819
Unsuccessful	383	0,8764	0,2638

Significance levels for variable entry and removal were fixed at the same levels as for first semester achievement. Allocation of observations to the two performance groups was excecuted as before.

5.3.4.3.4 Results

A summary of the classification results of the discriminant analysis procedure is given in table 5.11.

Table 5.11 Summary of discriminant analysis results for the 2-variable model.

DISCRIMINANT FUNCTION				
	Successful group		Unsuccessful group	
	Variables	Discriminant loading	Variables	Discriminant loading
	Swedcor Physics	0,3506 0,3194	Swedcor Physics	0,2726 0,2924
Percentage cases correctly classified	3,70 (2)		99,22 (380)	
Total error Rate†	96,30		0,78	

While performance levels in the Swedish rating and Physical science could therefore adequately predict potentially unsuccessful candidates, it is extremely limited in predicting potential success. The percentage of candidates correctly classified into the groups on the basis of the 2-variable model is marginally better than the corresponding values attained when using the 3-variable model. Should this model be adopted to select students onto the programs, the following results would have been attained for the given data.

Table 5.12 Final semester performance according to the 2-variable model

	Predicted Status			
		Success	**Failure**	**Total**
Actual Status	**Success**	**2**	**3**	**5**
	Failure	**52**	**380**	**432**
	Total	**54**	**383**	**437**

According to table 5.12, two students from the successful group and 52 students from the unsuccessful group would have been selected onto the course on the basis of the 2-variable model. Of the 54 students selected, the percentage of successful students in the group is 3,7% (2/54), while the successful students in the rejected group is 0,8% (3/383). We could correctly classify unsuccessful candidates with a much greater degree of certainty than we do successful candidates.

The probability of classifying possible successful candidates was in fact so low that none of the obtained models could be considered as significant.

5.3.4.4 Interpretation of the discriminant analyses

While both models are inadequate in correctly allocating students from the successful group, they have an exceptionally high success rate in correctly allocating students from the unsuccessful group.

The results indicate that, should the 3-variable model be applied in the selection process, we would expect:

None of the students admitted onto the course and approximately 12% of the rejected students would be successful. Stated differently, we would have made a correct decision in the selection of successful candidates in none of the cases.
However, this is based on results obtained from a sample consisting of a single observation, obtained through application of this model to a data set of 412 observations. In contrast, the success rate in correctly identifying unsuccessful candidates would be nearly 88%. The 3-variable model is thus not very effective in identifying potentially successful students. Should the 2-variable model be used as selection tool, we would expect the following results:

Approximately 4% of the accepted students and about 1% of the rejected students would be successful. Stated differently, we would have made a correct decision in the selection of successful candidates only 4% of the time. In contrast, the success rate in correctly identifying unsuccessful candidates would be 99%.

As was the case with the 3-variable model, the 2-variable model is also not very effective in identifying potentially successful students. Both models would however be effective in identifying potentially unsuccessfully students, with the 2-variable model performing better than the alternative 3-variable model. Thus none of the models provides satisfactory discriminant functions for the correct allocation of successful candidates to any of the two groups. The practical implications of these results are that:

1. Students rejected on the basis of these models would have been unsuccessful in their studies in excess of 88% of cases. Thus the correct decision in the identification of possible unsuccessful candidates would have been made in the majority of cases.
2. Students accepted on the ground of either of these models would have been unsuccessful in 96 or more percent of cases. A wrong decision in the identification of possible successful candidates would thus have been taken in the majority of cases. Such a situation would for obvious reasons be untenable for the institution.

3. Both models are expected to give rise to pass rates well below the current pass levels.

Chapter 6

Regression Analysis

6.1 Introduction

In this chapter we use regression analysis to examine the relationship between the dependent variable, student performance, and the set of independent variables.

6.2 Summary

The following observations were made regarding the students who wrote the aptitude test :

1. The models developed to describe success in individual subjects yielded better pass rates than those developed for the students who gained automatic entrance.

2. All the models had the Swedish rating as the single most important predictor of success.

3. Other variables identified as significant predictors of success for certain first semester subjects, were Mental alertness, and to a lesser extent Biology, Technical reading, Mechanical comprehension and Gottschaldt.

4. The percentage variation in individual subject performance explained by these models varied from 3,5% to 24%.

5. A Swedish rating of at least 35 and a nominal 45% pass level in the relevant additional predictor(s), would yield expected pass rates of at least 60% in most first semester subjects.

The following observations regarding the students who gained automatic entrance were made:

1. The models describing performance in individual first semester subjects showed, in all but one of the cases, the Swedish rating to be the only significant predictor. The only other significant variable was Biology.

2. The percentage variation in subject performance described by these models was 3% to 18%.

If we compare the results for the different subjects obtained through these models, we observe the following:

1. The expected pass rates for both sets of models indicated improved pass rates with improved Swedish ratings.

2. At most Swedish rating levels and for the majority of subjects, the models developed for the students who wrote the aptitude test consistently yielded better pass rates than those developed for the other group.

3. The improved R^2 values obtained for the group who wrote the aptitude test, compared with the other group, indicate that individual components in varying degrees improved our ability to define student success in specific first semester subjects.

6.3 Regression analysis

Regression analysis is used to describe the statistical relationship between a dependent variable Y and a set of independent (or explanatory) variables, say $X_1, X_2, \ldots, X_{p-1}$. The regression model is given by:

$$Y = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{p-1} X_{ip-1} + \varepsilon_i = \beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik} + \varepsilon_i$$

where the β_i's are unknown and called the regression coefficients and the ε_i's are statistical errors. The ε_i's ~ $N(0,\sigma^2)$, $i = 1,2,\ldots,n$. If $X_1 = X_2 = \ldots = X_{p-1} = 0$, then β_0 gives the mean response $E(Y)$ or $\hat{Y}$ at $X_1 = X_2 = \ldots = X_{p-1} = 0$, otherwise β_0 has no particular meaning as a separate term in the regression model.

The i-th regression coefficient β_i measures the influence of the X_i-th variable on Y with a unit increase in X_i while keeping all other independent variables constant. The values of the $\beta_1, \beta_2, \ldots, \beta_{p-1}$ are estimated from the data, giving rise to estimated values $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_{p-1}$. Given that $x_1, x_2, \ldots x_p$ are the observed values of $X_1, X_2, \ldots, X_p$ respectively, the predicted mean value of Y, $\hat{Y}$, is

$$E(Y) = \hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_{p-1} x_{p-1}.$$

An important statistic in regression analysis is the coefficient of determination, R^2. Let the multiple correlation coefficient between Y and the set of independent variables $X_1, X_2, \ldots, X_{p-1}$ be denoted by R. Then the proportion of variability in Y explained by $X_1, X_2, \ldots, X_{p-1}$ is denoted by R^2.

Introducing additional independent variables however does not always contribute in a statistically significant way towards explaining the variation in Y. The idea is to select a subset of independent variables that will best account for the variation in Y.

We attempt to explain the variation in Y as fully as possible by as few independent variables as possible. In this study we used the stepwise regression method to determine which variables should be included in the model.

Stepwise regression is a forward selection procedure that at each stage rechecks the importance of all variables previously included. If the partial sum of squares for any previously included variable does not meet the minimum criterion to stay in the model, the selection procedure changes to backward elimination and variables are individually dropped until those remaining meet the set criterion. Thereafter the forward selection procedure is resumed.

The criterion used for terminating the selection process is the ratio of the reduction in residual sum of squares caused by the next variable to be considered for inclusion to the residual mean square from the model inclusive of the variable. This criterion can be expressed as a critical "F to enter" or in terms of a critical "significance level to enter" where F is the F-test of the partial sum of squares of the variable being considered. The forward selection process is terminated when no further variables outside the model meet the minimum criterion to enter.

When high intercorrelation among the independent variables themselves exists, the estimated regression coefficients will have large sampling variability. This problem is commonly referred to as collinearity. In the event of collinearity amongst the independent variables, relatively small errors in Y would cause the regression coefficients to fluctuate considerably as variables are added to or removed from the model. This instability in the regression coefficients is manifested through very large standard errors for the partial regression coefficients and often none of these coefficients will be significantly different from zero even though their combined effect may be highly significant.

Hair *et al* (1991:37) commenting on the impact of collinearity on the predictor selection process proposes as a rule of thumb that; should a predictor be more closely related to the best predictor than it is to the dependent variable, it should not be entered into the model.

Another criterion of importance is Mallow's C_p statistic, which is concerned with the total mean squared error of the n fitted values for each of the various regression models. The model which includes all *p-1* potential independent variables is assumed to be chosen such that the mean squared error, $MSE(X_1, X_2, ..., X_{p-1})$ is an unbiased estimator of σ^2. When no bias occurs in the model with p - *1* independent variables, the expected C_p value is approximately p. Thus when the C_p values for all possible regression models are plotted against p, those models with little bias will tend to fall near the $C_p = p$ line, and vice versa.

In applying the C_p criterion one seeks to identify the subset of independent variables that have both a small C_p value and which is close to the value of p. In conclusion therefore, when C_p is small, total *MSE* is small and when C_p is close to p, bias in the model is small.

6.3.1 Regression analysis of first semester performance

Stepwise regression analysis was used as the model selection method in the identification of possible predictors. As a precursor to the analysis, a correlation matrix of all variables, dependent and independent, was obtained to ascertain possible collinearities among variables. A summary of the correlation matrix is given in table 6.1. All first semester subjects were then individually regressed on a combination of the variables obtained during Chaid, discriminant and regression analyses.

Variables used as predictors in the stepwise regression analysis process were:

Dependent variables: Engineering Mathematics 1 (Engmat1), Digital Systems 1 (Digsy1), Electronics 1 (Elect1), Electrical Engineering1 (Elen1), Industrial Technology 1 (Intec1), Engineering Mechanics 1 (Engme1), Communication Languages (Comla), Communication Programming1 (Compr1) and Machine Drawings1 (Macdr1).

Predictor variables: Swedish rating (Swedcor), Matric aggregate (Agg), Physical Science (Phys), Biology, Mathematics (Maths), Afrikaans (Afr), English (Eng), Mental ability (Men), Mechanical comprehension (Mech), Technical reading ability (Read), Gottschaldt figure test (Gots) and aptitude test total (Total).

Table 6.1 displays the correlations among the independent variables, and with the dependent variables. Examination of the correlation matrix indicates that the Swedish rating is more closely related to the independent variables than any other predictor, ranging between 0,18 and 0,44. Other predictors closely following Swedish rating are matric aggregate, Biology and Mathematics. A final observation concerns the low to moderate correlations between the predictors and the dependent variables, while those predictors more highly correlated with the dependent variables also have relatively high correlations among each other.

When applying the rule of thumb regarding collinearity proposed by Hair *et al*, the predictors matric aggregate, Mathematics, Biology, Physics, Afrikaans and English would probably not be appropriate for the model building process. This follows since most of the dependent variables are more highly correlated with Swedish rating than with the respective dependent variables. All such variables were, where applicable, excluded from development of appropriate regression equations.

Table 6.1 Summary of correlation matrix

Dependent Variable	Predictor Variable											
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
Y_1 = ENGMAT1	0,39	0,23	0,15	0,21	0,21	0,07	0,11	0,09	0,05	-0,01	0,15	0,09
Y_2 = DIGSY1	0,29	0,23	0,13	0,22	0,14	0,11	0,19	0,17	0,10	0,03	0,11	0,15
Y_3 = ELECT1	0,28	0,21	0,17	0,21	0,15	0,03	0,12	0,16	0,04	0,05	0,15	0,11
Y_4 = ELEN1	0,36	0,22	0,18	0,21	0,20	0,09	0,15	0,07	0,09	-0,02	0,12	0,07
Y_5 = INTEC1	0,18	0,14	0,08	0,10	0,10	0,07	0,18	-0,03	-0,02	-0,08	0,11	0,00
Y_6 = ENGME	0,44	0,28	0,22	0,19	0,25	0,18	0,24	0,12	0,15	-0,03	0,19	0,15
Y_7 = COMLA	0,19	0,17	0,07	0,16	0,10	0,12	0,16	0,11	0,04	-0,03	0,09	0,06
Y_8 = COMPR1	0,36	0,32	0,20	0,22	0,16	0,17	0,23	0,15	0,11	0,03	0,12	0,11
Y_9 = MACDR	0,35	0,32	0,16	0,19	0,16	0,26	0,23	0,23	0,22	-0,03	0,26	0,20
Predictor Variable	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
X_1 = SWEDCOR	1,00											
X_2 = AGG	0,69	1,00										
X_3 = PHYSICS	0,54	0,45	1,00									
X_4 = BIOLOGY	0,28	0,58	0,28	1,00								
X_5 = MATHS	0,53	0,42	0,34	0,28	1,00							
X_6 = AFR	0,41	0,45	0,12	0,34	0,10	1,00						
X_7 = ENG	0,41	0,43	0,14	0,16	0,04	0,36	1,00					
X_8 = MEN	0,27	0,21	0,01	0,15	0,11	0,33	0,16	1,00				
X_9 = MECH	0,18	0,14	0,05	0,13	0,02	0,28	0,12	0,46	1,00			
X_{10} = READ	0,07	0,09	-0,00	0,10	0,04	0,13	0,10	0,33	0,32	1,00		
X_{11} = GOTS	0,22	0,16	0,06	0,13	0,01	0,27	0,13	0,46	0,40	0,19	1,00	
X_{12} = TOTAL	0,23	0,19	0,02	0,16	0,06	0,34	0,18	0,75	0,69	0,58	0,71	1,00

The correlation coefficients of both Swedish rating and English with Digital Systems were similar. Each predictor was individually used to determine an appropriate regression model for Digital Systems.

The independent variables were subsequently regressed on the identified set of predictor variables using the SAS PROC REG procedure. The forward selection procedure of stepwise regression was used to determine which variables should be included in the model. The probability for variables to either enter or stay in the model was set at 0,15. A summary of the models developed through this process is displayed in table 6.2. These results offered no surprises as Swedish rating emerged as the most important predictor in all but a few of the models.

The predicted (estimated) value of each independent variable at different levels of the predictor variables could then be calculated by adding the intercept value to the sum of the product of regression coefficients (column 2) and the value of the predictor variables (column 1). So, for example, the regression estimate for Engmat 1 under Option 1 can be indicated by:

Engmat 1 = 6,71 + 1,53 Swedcor.

Where more than one predictor is identified as significant, the standardised regression coefficients or $\beta - coefficients$ (column 4), can be used to compare the relative importance of individual independent variables in relation to the indicated dependent variables. Consider for example the regression equation developed for Digital Systems under option 1. Here the Swedish rating ($\beta = 0,30$) is more influential in describing performance in Digital Systems than is Mental alertness ($\beta = 0,14$).

The standard error of the parameter estimate (column 3) should preferably be as small as possible, as smaller standard errors indicate more reliable estimates. The partial t-values indicate that all included variables are statistically significant predictors of the respective dependent variables. While a few of the predictors had a negative effect on student performance in the different subjects (indicated by a negative sign associated with the particular predictor's regression coefficient), most had positive effects. An improvement in performance of those predictors with positive regression coefficient is expected to improve student performance in individual first semester subjects. The magnitude of this increase would depend on the strength of their respective regression coefficients. In the case of negative regression coefficients, increases in the independent variables are expected to negatively influence the students' performance in the affected subjects, in accordance with the size of the regression coefficient. Further information on these models is displayed in table 6.3.

The matric variables included as predictors of first semester success have a positive impact on performance in these subjects. The aptitude test components however have differing effects on different first semester subjects. For example, Technical reading ability has a negative impact on a student's performance in Communication but a positive influence on performance in Machine drawings.

Whenever Swedish rating was included in the stepwise regression procedure it emerged as an important predictor. Other important predictors in option 1 were Biology, English, Mental alertness, Mechanical comprehension, Technical reading ability and Gottschaldt figure test. English, Biology and Swedish rating emerged as important predictors in option 2.

Table 6.2 Regression models developed for different independent variables

Independent variable	Included predictors	Regression coefficients	Standard error of coefficients [se(b)]	Standard regression coefficients	Partial t value
OPTION 1 (IF BOTH MATRIC AND APTITUDE RESULTS ARE AVAILABLE)					
ENGMAT1	INTERCEPT	6,71	5,80		
	SWEDCOR	1,53	0,20	0,43	60,50
DIGSY1	INTERCEPT	2,81	10,18		
	SWEDCOR	1,50	0,29	0,30	16,70
	MEN	2,15	1,30	0,14	2,73
ELECT1	INTERCEPT	-0,65	10,17		
	SWEDCOR	1,14	0,28	0,32	19,16
	MEN	1,98	1,30	0,12	2,35
ELEN1	INTERCEPT	-5,97	8,43		
	SWEDCOR	1,47	0,27	0,37	29,14
INTEC1	INTERCEPT	35,95	7,35		
	SWEDCOR	0,72	0,22	0,23	9,24
	MEN	-2,86	1,26	-0,17	2,18
	GOTS	2,41	1,12	0,16	4,62
INTEC1	INTERCEPT	25,26	10,22		
	ENG	0,51	0,19	0,19	6,91
ENGME1	INTERCEPT	5,09	6,70		
	SWEDCOR	1,60	0,26	0,47	38,70
COMLA	INTERCEPT	32,45	7,30		
	SWEDCOR	0,41	0,19	0,15	13,03
	BIOLOGY	0,24	0,12	0,14	3,61
	READ	-1,03	0,66	-0,09	2,43
COMPR1	INTERCEPT	24,87	5,20		
	SWEDCOR	1,09	0,18	0,36	38,16
MACDR1	INTERCEPT	6,22	10,16		
	SWEDCOR	1,01	0,24	0,34	18,92
	MECH	-3,23	1,10	-0,28	6,67
	READ	6,96	1,82	0,36	8,70
OPTION 2 (IF ONLY MATRIC RESULTS ARE AVAILABLE)					
ENGMAT1	INTERCEPT	12,30	4,73		
	SWEDCOR	1,35	0,16	0,40	73,79
DIGSY1	INTERCEPT	20,74	6,29		
	SWEDCOR	0,89	0,20	0,29	20,08
ELECT1	INTERCEPT	19,17	6,37		
	SWEDCOR	0,88	0,20	0,28	19,67
ELEN1	INTERCEPT	1,44	6,20		
	SWEDCOR	1,29	0,20	0,36	43,12
INTEC1	INTERCEPT	36,66	4,88		
	SWEDCOR	0,53	0,16	0,17	10,21
INTEC1	INTERCEPT	23,80	7,66		
	ENG	0,54	0,15	0,21	13,55
ENGME1	INTERCEPT	6,80	6,05		
	SWEDCOR	1,46	0,22	0,42	45,91
COMLA	INTERCEPT	30,22	4,99		
	SWEDCOR	0,31	0,15	0,12	13,58
	BIOLOGY	0,20	0,10	0,11	3,97
COMPR1	INTERCEPT	25,86	4,19		
	SWEDCOR	1,02	0,14	0,34	53,34
MACDR1	INTERCEPT	16,78	6,12		
	SWEDCOR	0,98	0,22	0,32	19,45

Where Swedish rating emerged as an important predictor, a comparison of standardised $\beta - values$ indicated that it was the most influential predictor in all but one of the regression functions. For option 1 combinations, the percentage of explained variation in individual first semester subject performance varied between 3,5% for Industrial Technology to 24% for Machine drawings 1. The respective lowest and highest R^2-values for option 2 were 3% and 18% respectively. The difference in the explained variation between the two options is depicted in figure 6.4.

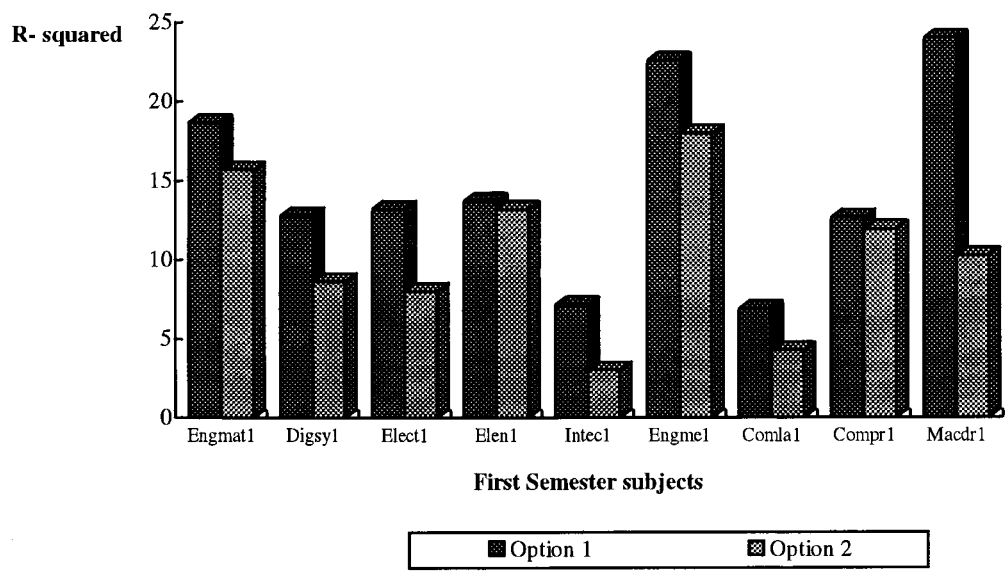
Table 6.3 Summary of the most important school and aptitude test explanatory variables for the listed first semester subjects

Dependent Variables	Predictors	Partial R^2	R^2	C_p	p	F	Prob > F
OPTION 1							
ENGMAT1	SWEDCOR	0,1865	0,1865	-0,52	2	60,50	0,000
DIGSY1	SWEDCOR	0,1101	0,1101	-0,53	2	16,70	0,000
	MEN	0,0178	0,1279	-1,18	3	2,73	0,100
ELECT1	SWEDCOR	0,1184	0,1184	2,65	2	19,16	0,000
	MEN	0,0143	0,1325	2,31	3	2,35	0,127
ELEN1	SWEDCOR	0,1374	0,1374	-0,95	2	29,14	0,000
INTEC1	SWEDCOR	0,0416	0,0416	3,97	2	9,24	0,002
	MEN	0,0097	0,0513	3,78	3	2,18	0,141
	GOTS	0,0203	0,0717	1,22	4	4,62	0,032
INTEC1	ENG	0,0347	0,0347	0,06	2	6,91	0,009
ENGME1	SWEDCOR	0,2254	0,2254	4,03	2	38,70	0,000
COMLA	SWEDCOR	0,0474	0,0474	4,62	2	13,03	0,000
	BIOLOGY	0,0130	0,0604	3,01	3	3,61	0,058
	READ	0,0087	0,0691	2,59	4	2,43	0,120
COMPR1	SWEDCOR	0,1267	0,1267	2,32	2	38,16	0,000
MACDR1	SWEDCOR	0,1374	0,1374	12,56	2	18,92	0,000
	MECH	0,0462	0,1833	7,63	3	6,67	0,011
	READ	0,0565	0,2398	1,15	4	8,70	0,003
OPTION 2							
ENGMAT1	SWEDCOR	0,1571	0,1571	1,10	2	73,79	0,000
DIGSY1	SWEDCOR	0,0865	0,0865	2,60	2	20,08	0,000
ELECT1	SWEDCOR	0,0804	0,0804	2,29	2	19,67	0,000
ELEN1	SWEDCOR	0,1322	0,1322	1,17	2	43,12	0,000
INTEC1	SWEDCOR	0,0302	0,0302	1,13	2	10.21	0,001
INTEC1	ENG	0,0438	0,0438	-0,01	2	13,55	0,000
ENGME1	SWEDCOR	0,1801	0,1801	1,05	2	45,91	0,000
COMLA	SWEDCOR	0,0332	0,0332	4,97	2	13,58	0,000
	BIOLOGY	0,0096	0,0429	3,00	3	3,97	0,047
COMPR1	SWEDCOR	0,1197	0,1197	2,20	2	53,34	0,000
MACDR1	SWEDCOR	0,1027	0,1027	1,93	2	19,45	0,000

The information contained in table 6.3 indicates similar regression models for Engineering Mathematics, Electrical Engineering, Engineering Mechanics and Communication Programming irrespective of the student group the data derived from. Although the R^2 values from option 1 models were always better than those of option 2 models, these differences were not very high. However, the R^2 differences for the rest of the first semester subjects were more pronounced. These differences were mainly due to the additive effect that the inclusion of aptitude test variables into the regression models had on the percentage explained variation. This is an indication of the importance of these variables in explaining success in these subjects.

Generally, a low percentage of the variation in performance of Technikon subjects could be accounted for by the various regression models.

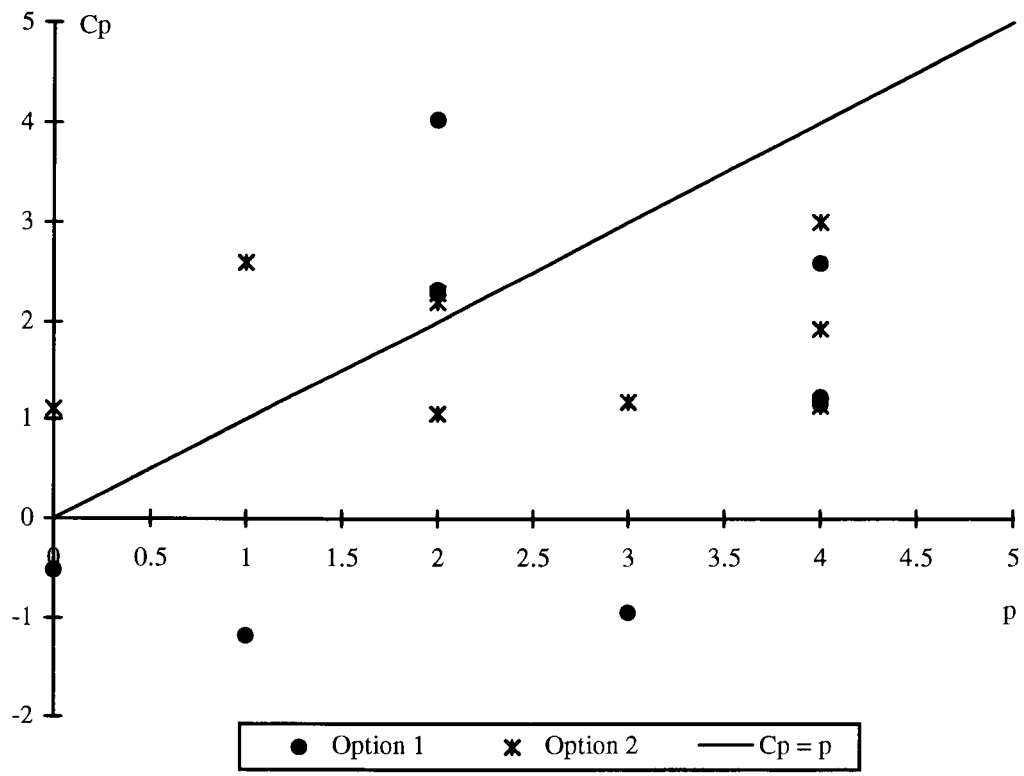
Figure 6.4 Comparison of R^2 values obtained for the regression models used in option 1 and option 2



All predictors included were found to be statistically significant (refer F and Prob >F-values) in table 6.3.

If we compare the Mallows C_p statistic (column 5) with the p-values (column 6) for the specific models, i.e. the number of significant explanatory variables in the resultant model for each combination plus the independent variable, the following becomes clear: C_p values for both options 1 and 2 were generally lower than their corresponding p-values. Figure 6.5 depicts the deviation of the Mallows C_p values from the p-values for both options 1 and 2. This figure indicates that the model obtained for option 2 is more closely distributed around the C_p = p line than that obtained for option 1.

Figure 6.5 Graph of Cp versus p



6.3.2 Expected first semester pass rates

The assumptions ε_i *iid* $N(0,\sigma^2)$ and $E(Y/x_1,x_2,\ldots,x_{p-1}) = \beta_0 + \sum_{i=1}^{p-1}\beta_i x_i$ imply that the conditional distribution of Y given $x_1,x_2,\ldots,x_{p-1}$ is $Y/x_1,x_2,\ldots,x_{p-1} \sim N(\beta_0 + \sum_{i=1}^{p-1}\beta_i x_i, \sigma^2)$.

Given this conditional distribution the pass rate, $p(x_1, x_2, \ldots, x_{p-1})$, is calculated as

$$p(x_1, x_2, \ldots, x_{p-1}) = \int_{50}^{100} n(y, \beta_0 + \sum_{i=1}^{p-1} \beta_i x_i, \sigma^2) dy$$

where $n(y, \mu, \sigma^2)$ represents the normal density with mean μ and variance σ^2.

The expected pass rates at different Swedish rating levels, assuming a pass mark of 50% in individual first semester subjects, were then calculated for the two models. These pass rates are displayed in table 6.6.

This table depicts the expected pass rates in individual first semester subjects for Swedish rating levels ranging from 25 to 45. Where predictors other than Swedish rating existed, the required levels of competence for all such predictors were fixed at 45%. This is of course an arbitrary choice since a range of possible entry level combinations for these predictors could be used.

Column 2 of table 6.6 indicates the explanatory variables found to be significant in describing the variation in the dependent variables listed in column 1, while the expected pass rates at different performance levels are given in columns 3 to 7.

Table 6.6 Expected pass rates in individual first semester subjects

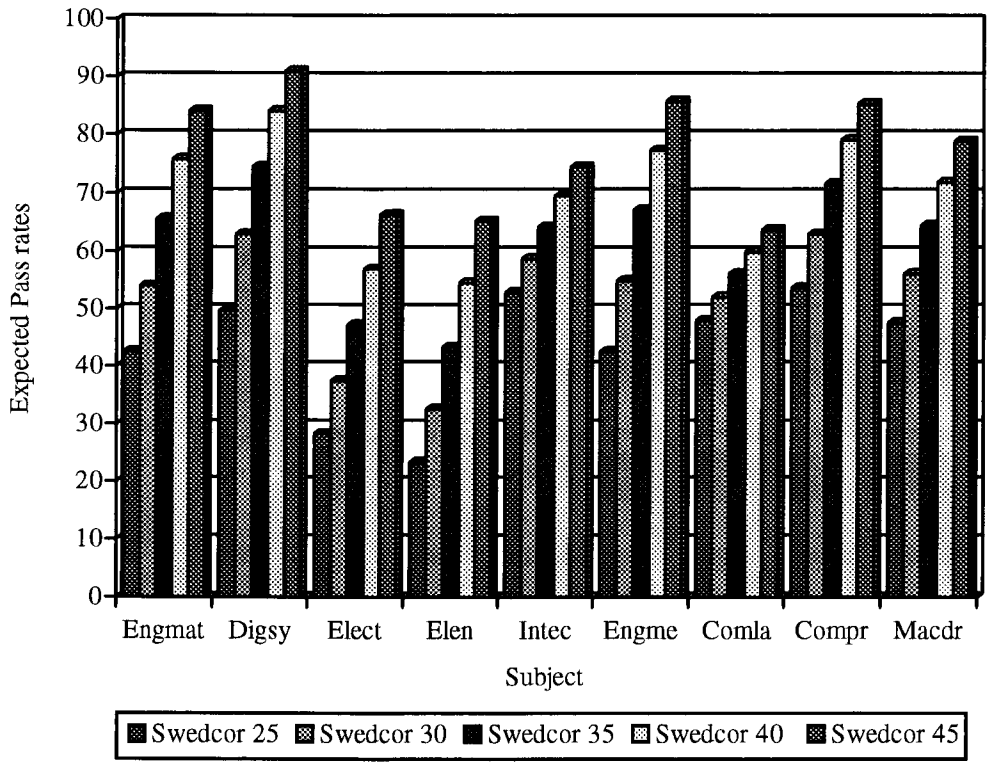
First semester subjects	Predictor variables	Swedish rating levels				
		25	30	35	40	45
OPTION 1						
Engineering mathematics	Swedcor	42,5	54,0	65,5	75,8	84,1
Digital Systems	Swedcor, Men	49,6	62,9	74,5	84,1	91,0
Electronics	Swedcor, Men	28,4	37,5	47,2	56,8	66,3
Electrical Engineering	Swedcor	23,3	32,6	43,3	54,4	65,2
Industrial technology	Swedcor, Men, Gots	52,8	58,7	64,1	69,5	74,5
Engineering mechanics	Swedcor	42,5	54,8	67,0	77,3	85,8
Communication languages	Swedcor, Biology, Read	48,0	52,0	56,0	59,9	63,7
Communication programming	Swedcor	53,6	62,9	71,6	79,1	85,3
Machine drawings	Swedcor, Mech, Read	47,6	56,0	64,4	71,9	78,8
OPTION 2						
Engineering mathematics	Swedcor	43,6	54,4	64,8	74,2	82,1
Digital Systems	Swedcor	37,8	45,6	53,2	61,0	68,4
Electronics	Swedcor	35,2	42,5	50,0	57,5	64,8
Electrical Engineering	Swedcor	27,1	35,6	44,8	54,4	64,1
Industrial technology	Swedcor	50,0	54,4	58,3	62,6	66,6
Engineering mechanics	Swedcor	39,0	51,2	62,9	73,9	82,6
Communication languages	Swedcor, Biology	48,0	50,8	54,0	57,1	59,9
Communication programming	Swedcor	52,4	61,4	69,9	77,0	83,4
Machine drawings	Swedcor	34,1	42,9	52,0	61,0	69,5

Figure 6.7 depicts the expected pass rates at different Swedish ratings for the students who wrote the aptitude test, while figure 6.8 represents the pass rates for those who gained automatic entrance.

If the Swedish rating level is fixed at 35, those students who wrote the aptitude test are expected to obtain pass levels above 60% in all subjects except Electrical Engineering (43%), Communication Languages (56%) and Electronics (47,2%). In order to raise the pass rates in these three subjects to comparable levels through manipulation of the Swedish rating scores, one would require Swedish rating levels in excess of approximately 45. Raising the Swedish rating entrance level to 45 may create problems in terms of accessibilty to Higher education for some students and hence might be impractical.

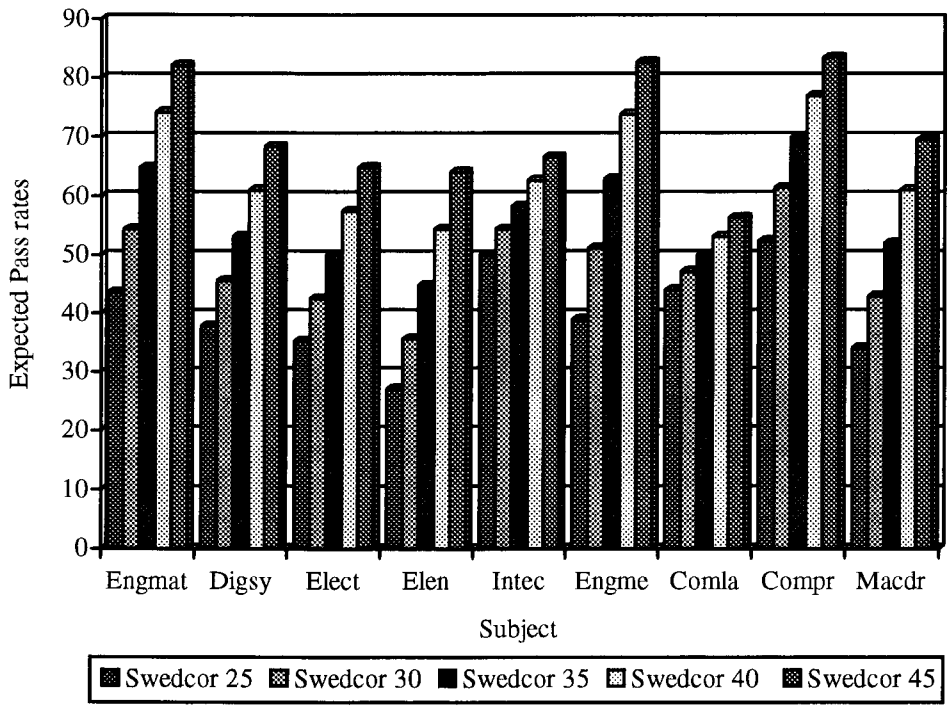
Serious consideration should however be given to an entrance level of 35.

Figure 6.7 Graph of expected pass rates at different Swedish rating levels (Option 1)



The performance of the students who gained automatic entrance (refer figure 6.8) indicates that these students would require higher Swedish rating levels than their option 1 counterparts in order to attain similar pass rates. For example, at a Swedish rating level of 35, pass rates in excess of 60% would only be possible for three out of the nine subjects.

Figure 6.8 Graph of expected pass rates at different Swedish rating levels (Option 2)



Since Swedish rating and other predictor variable levels are obviously a function of the pass level requirements, set according to School objectives, setting appropriate and acceptable entry levels of competence in line with these objectives, are extremely important. The graph in figures 6.7 and 6.8 (or variations thereof) could be used to assist in this process.

The most obvious observation is the increase in subject pass rate with an increase in Swedish rating level. A comparative analysis of expected performance in individual subjects for the two options under consideration, is depicted in figures 6.9 to 6.17.

Figure 6.9 Comparison of performance in Engineering Mathematics 1 for Options 1 & 2 at different Swedish rating levels

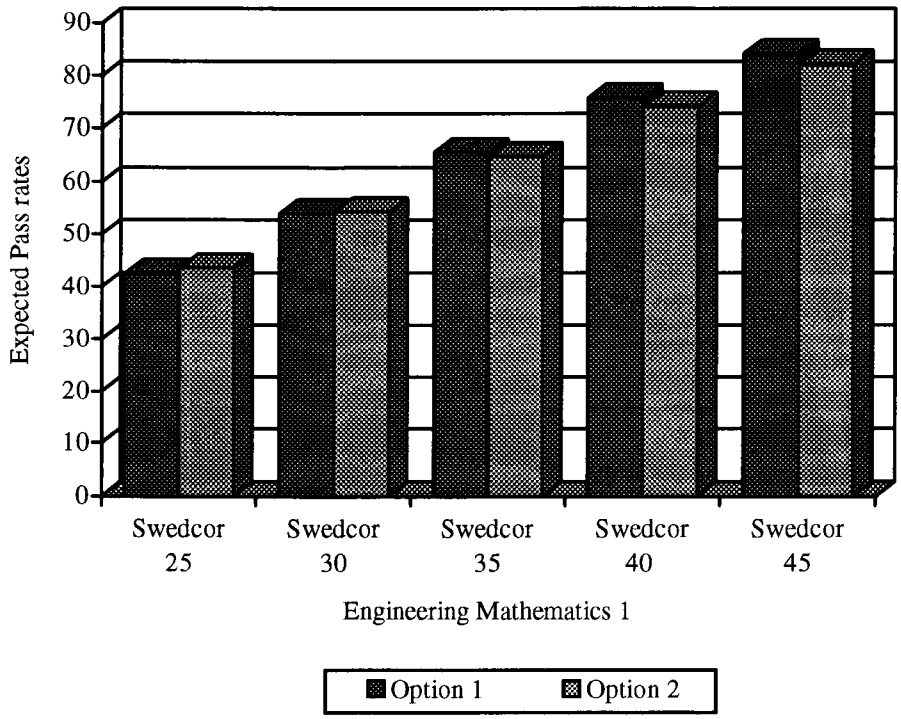


Figure 6.10 Comparison of performance in Digital Systems 1 for Options 1 & 2 at different Swedish rating levels

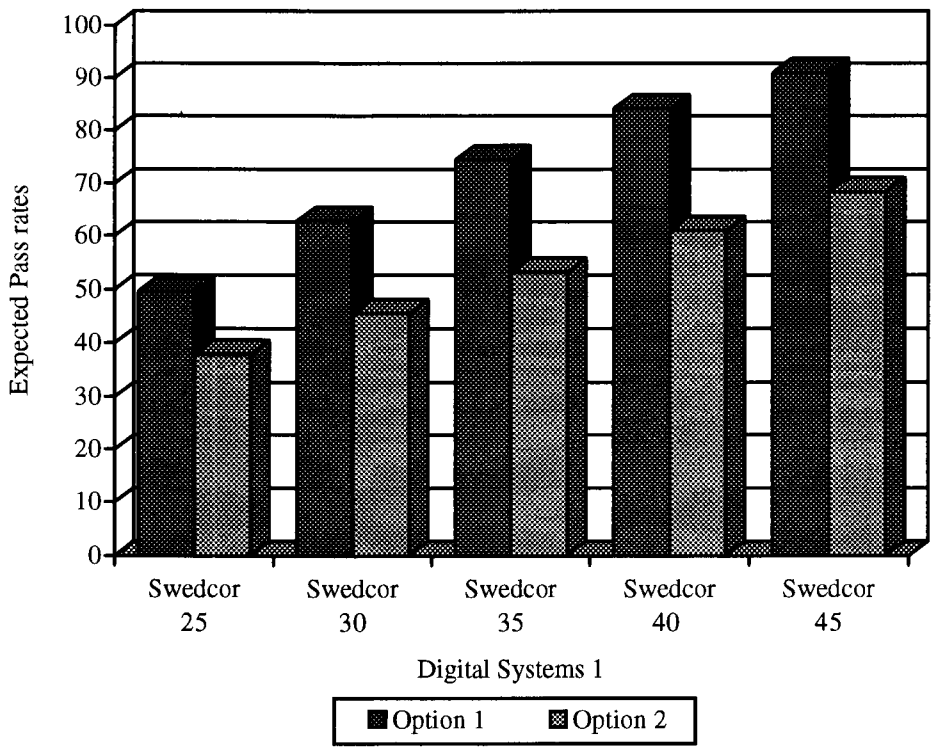


Figure 6.11 Comparison of performance in Electronics 1 for Options 1 & 2 at different Swedish rating levels

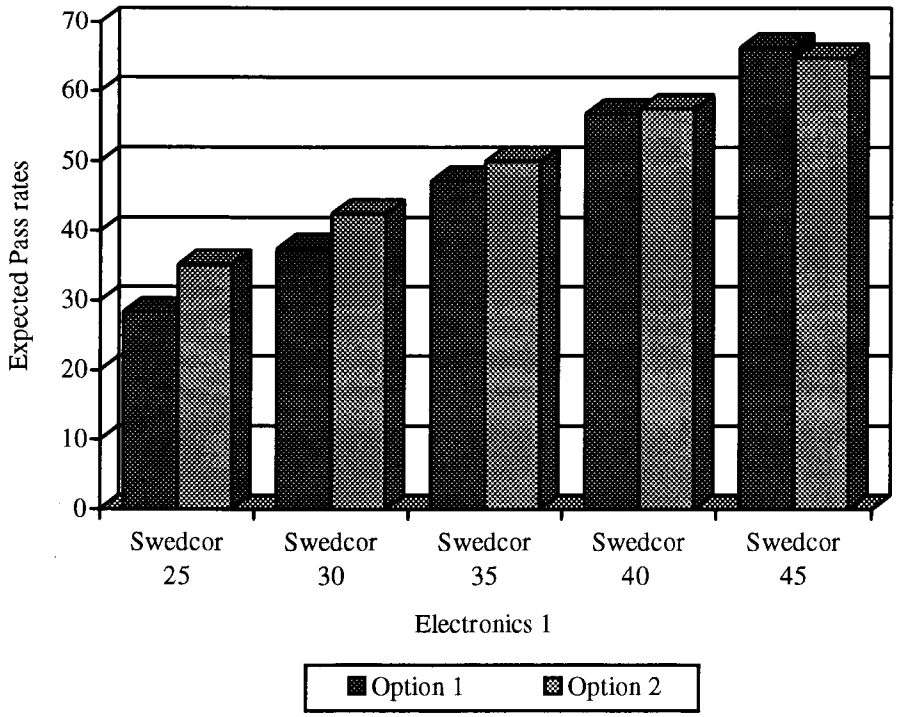


Figure 6.12 Comparison of performance in Electrical Engineering 1 for Options 1 & 2 at different Swedish rating levels

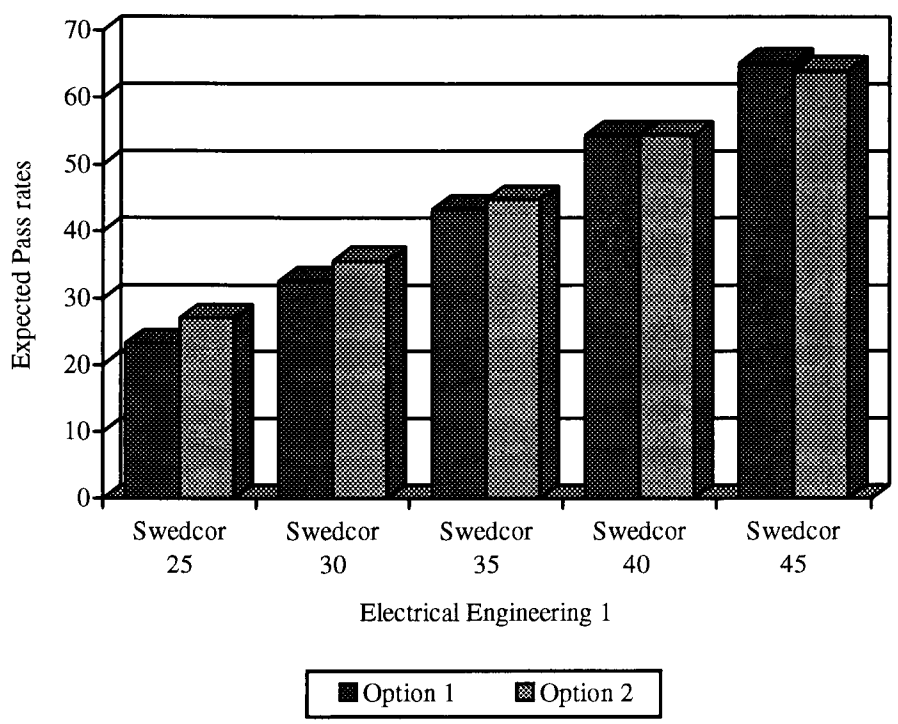


Figure 6.13 Comparison of performance in Industrial Technology 1 for Options 1 & 2 at different Swedish rating levels

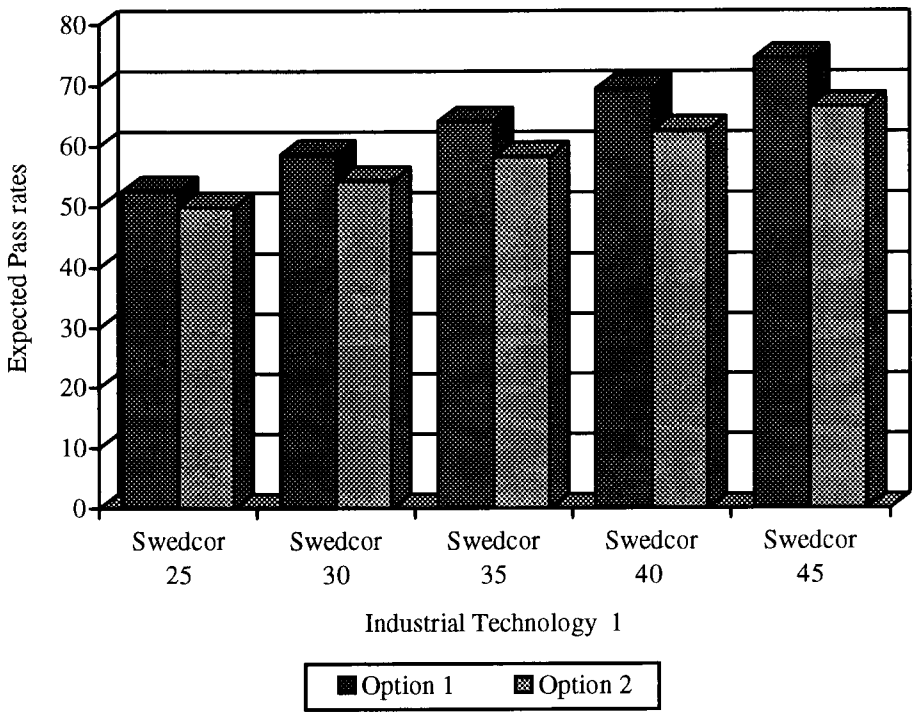


Figure 6.14 Comparison of performance in Engineering Mechanics 1 for Options 1 & 2 at different Swedish rating levels

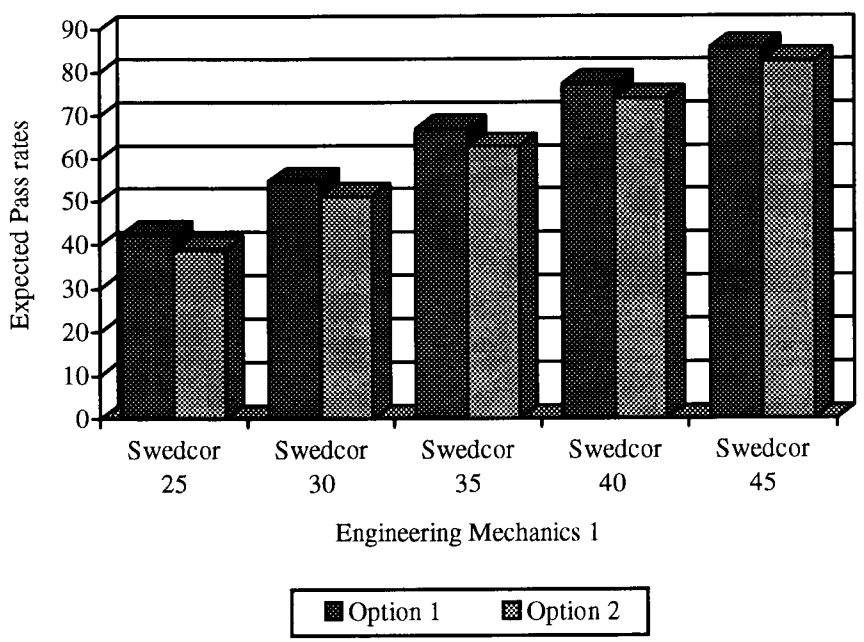


Figure 6.15 Comparison of performance in Communication Languages for Options 1 & 2 at different Swedish rating levels

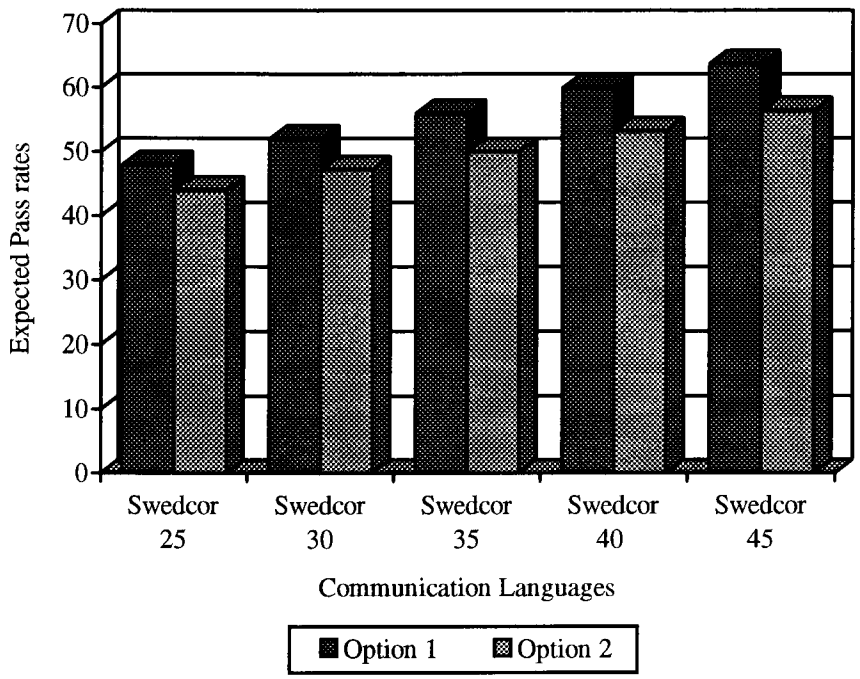


Figure 6.16 Comparison of performance in Communication Programming 1 for Options 1 & 2 at different Swedish rating levels

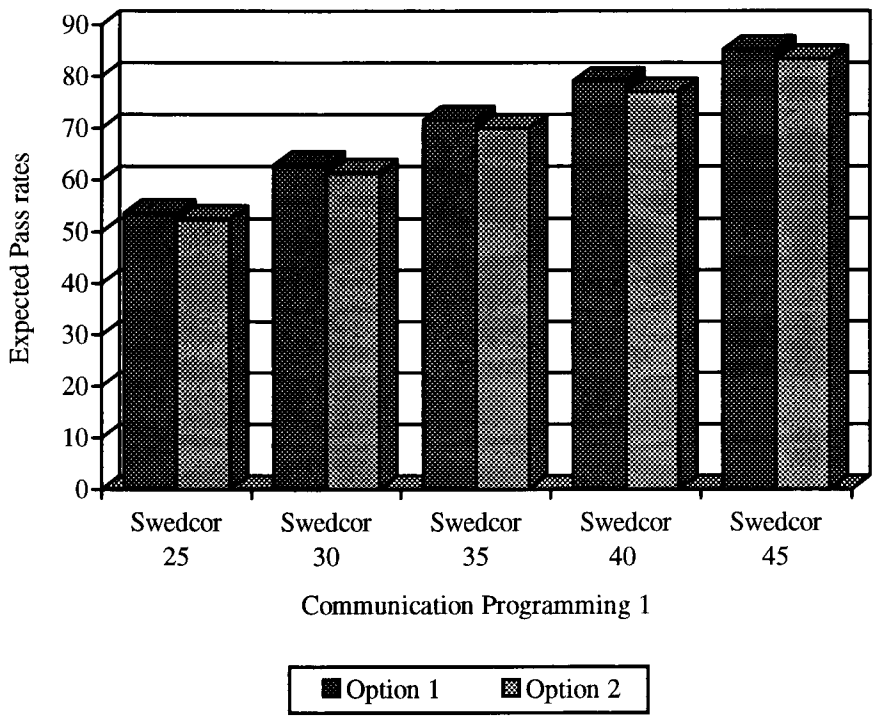
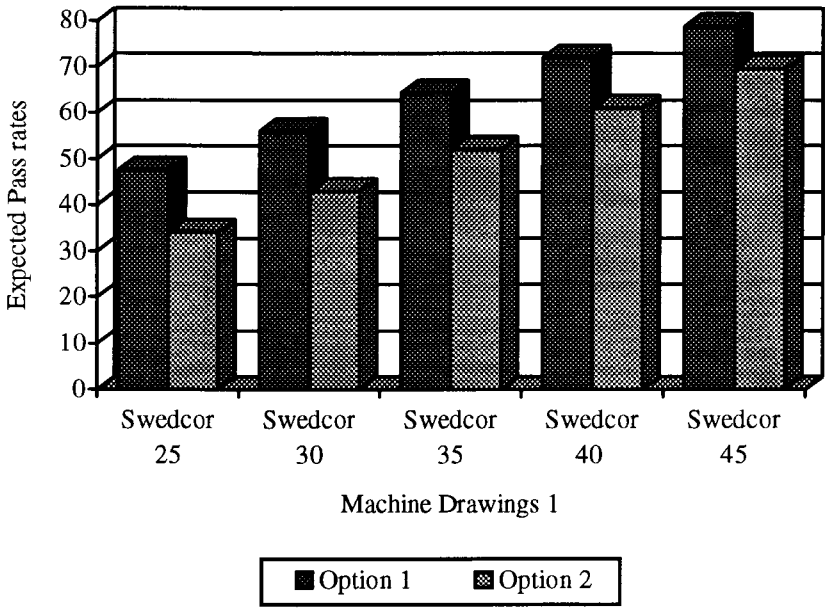


Figure 6.17 Comparison of performance in Machine Drawings 1 for Options 1 & 2 at different Swedish rating levels



Comparing expected pass rates for the two options, the following deductions were made:

For Engineering Mathematics comparative expected pass levels for the options would not differ considerably. Also, at the lower end of the Swedish rating scale, option 1 values exceeded comparative option 2 values. This trend was reversed at the upper end of the scale.

Comparative option 1 and 2 values for Digital system performance were considerably different, with option 1 values larger than their counterparts. The observed differences were more pronounced at the upper end of the Swedish rating scale.

The inter-option differences for both Electronics and Electrical Engineering were insignificant. Furthermore, while option 2 values were greater at the lower level, the trend was reversed at the upper end.

Although inter-option differences for Industrial Technology were relatively small, option 1 values, at improved Swedish rating levels, were increasing at a slightly higher rate than did their counterparts.

There was a pronounced improvement in Engineering Mechanics performance with improved Swedish rating levels for both options. Although option 2 values were consistently higher than comparative option 1 values, the inter-option differences were however not large.

Although the gradient of expected pass rate improvement was relatively flat for Communication Languages there was a definite improvement in pass rate with improved Swedish ratings.

Although option 1 values in Communication Programming were consistently higher than comparative option 2 values, the inter-option differences remained negligible throughout.

The inter-option differences in Machine Drawing performance at differing Swedish rating levels were relatively large. Expected pass rates at option 1 level exceeded those at comparative option 2 levels. Again, the general tendency was that of improved pass rates with improved Swedish ratings.

6.4 Conclusions

Results obtained in this chapter reinforced the standing of the Swedish rating as by far the most important predictor of student success at the first semester level. Improved Swedish rating performance invariably lead to varying degrees of improved performance in individual subjects. Furthermore, individual aptitude test components significantly improved our ability to model student success in certain first semester subjects.

Chapter 7

Conclusions and Recommendations

7.1 Conclusions

On the strength of the results from the different analyses performed, the following conclusions were reached:

7.1.1 Student selection models

Two separate models to assist in the discrimination between potentially successful or unsuccessful students were developed. We developed a model firstly, for the students who wrote the aptitude test, and secondly one for those who gained automatic entrance.

The best model developed to identify potential success on the part of the students who wrote the aptitude test (the 'aptitude test' group) was the 6-variable model, consisting of the Swedish rating, Technical reading, Mechanical comprehension, Afrikaans, Mathematics and English. This model would yield an expected overall pass rate of 68% as opposed to approximately 46% with the current models.

All the regression models developed to define probable success in individual first semester subjects for the 'aptitude test' group included the Swedish rating as the principal predictor of success.

However, if we use the 3-variable model developed for the students that gained automatic entrance, the expected overall pass rate would be approximately 60%. This model consisted of the Swedish rating, Mathematics and Afrikaans.

All but one of the regression models developed to define probable success in individual first semester subjects consisted solely of the Swedish rating.

7.1.2 School study record and personal characteristics

According to the Chaid and Discriminant analyses, the Swedish rating and the students' activity before first registration at Peninsula Technikon were the most important predictors of success in the first semester.

Furthermore, predictors identified as important in determining success at final semester level were Swedish rating, activity before first registration at Peninsula Technikon and Biology.

7.1.3 Swedish rating

Although there is strong evidence that the Swedish rating had to be adjusted, the research found no meaningful improvement in the pass level at the revised level of 32, initiated in 1992. However, a considerable percentage of students would be denied access to higher education by any increase in the Swedish rating cut-off point.

7.1.4 Pass rates

Pass rates at both first semester and full course levels were unacceptably low.

7.1.5 Aptitude test results

The significance of the aptitude test total in the identification of potentially successful candidates could not be established. This measure is therefore inadequate in identifying potentially successful candidates. A possible reason for this phenomenon is that the aptitude test results are primarily obtained for those students at the bottom end of the range. The situation could have been quite different if these results had been available for all students.

Individual aptitude test components were however found to be significant in assisting to identify possible success in certain first semester subjects. Acceptable levels of performance in first semester subjects could be established through manipulation of the Swedish rating scores and performance levels in particular aptitude test components.

7.1.6 Home Language

Afrikaans speaking students have a much greater probability of failure than other students. This could possibly be because tuition at the institution is provided in English.

7.1.7 Other factors

Distinct groupings of students at risk were identified. These included students from the HOR, those with a practical matric and students entering the Technikon immediately after completing their secondary education. Those with a DET qualification, matric with a full exemption and students with previous work experience performed considerably better than those without these qualifications. Furthermore, students who had previously registered at other tertiary institutions performed better than those without such experience.

7.2 Recommendations

7.2.1 Predictors

A more comprehensive range of potential predictors, possibly covering factors such as socio-economic, motivational, study-environment, study methods and habits, etc., needs to be included in any future study to find a more representative list of predictors that could be used to identify potentially successful students.

7.2.2 Pass rates and accessibility

To improve the pass rates in first semester subjects, the Swedish rating cut-off point should be set at 35. Should the need exist to further improve the pass rate without further increasing the Swedish rating level, this could be achieved through manipulation of the performance levels in the aptitude test components.

The increased Swedish rating cut-off point furthermore has to be accompanied by the introduction of well-constructed access programs. These programs would be important to counter the negative impact the increase would necessarily have on the accessibility to higher education for disadvantaged students.

It is also necessary to identify and implement the necessary academic intervention programs to redress the current situation of appallingly low pass rates in the full course.

7.2.3 Individual matric subjects

The emergence of Biology as an important factor in distinguishing between potentially successful and unsuccessful students needs to be investigated.

7.2.4 Aptitude test

A critical review of the use of the aptitude test in its current form is necessary. Particular attention should be given to the composition of the statistic to be used in any future selection processes.

7.2.5 Other

Those categories of students identified as possible poor performers have to be researched and the necessary remedial action taken.

7.3 Suggestions for further research

More effort is needed to develop the required levels of academic competence, motivation and aptitude necessary for successful study, in terms of either the existing criteria or ones specifically developed for the diverse student population.

The following issues should provide a challenge to future researchers in this field:

Can we find predictors that will significantly improve R^2 ?

Appropriate aptitude and motivation tests and norms need to be developed for the diverse groups being catered for.

Further tests should be performed to make the necessary changes, if any, to the Swedish rating or a unique measure could be developed that would adequately differentiate between potentially successful and unsuccessful students. Coupled with this is the establishment of appropriate factor weightings for the different components of a revised Swedish rating or any other appropriate measure that might be developed.

Bibliography

DRAPER N.R. & SMITH H., (1981), *Applied regression analysis*, Second edition. John Wiley & Sons Inc., New York.

DU TOIT S.H.C., STEYN A.G.W. & STUMPF R.H., (1984), *Statistical Graphics*, HSRC Report WS-30, HSRC, Pretoria.

GOUWS D.J., (1961), *Die Akademiese Vordering en aanpassing van Eerstejaar Universiteit studente,* Van Schaik Bpk., Pretoria

HAIR J.J., ANDERSON R.E., TATHAM R.L. & BLACK W.C., (1992), *Multivariate data analysis with readings*, Third edition. Macmillan Publishing Company, New York, N.Y.

INGRAHAM K. R. et al., *1988 SAT/STAT User Guide*, Release 6.03 Edition, SAS Institute Inc., NC, USA

JOHNSON R.A. & WICHERN D.W., (1988), *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey

KENDALL M., (1975), *Multivariate analysis*, Charles Griffin & Company Ltd., London.

KLEINBAUN D.G., KUPPER L.L. & MULLER K.E. (Second Edition), (1988), *Applied Regression Analysis and Other Multivariate Methods*, PWS - Kent Publishing Company, Boston, Massaschusetts

LEE I. & MINAKO K.M., (1991), *SAS Manual for Prentice Hall Statistics*, Prentice-Hall Inc., Englewood Cliffs, New Jersey.

MALHERBE E.G. & COOK P.A.W., (1938), *The Relationship of Entrance Age of University Students to their Academic Success*, Van Schaik Ltd, Pretoria

MANLY B.F.J.,(1986), *Multivariate statistical methods. A primer*, Chapman & Hall. New York, N.Y.

MARDIA K.V., KENT J.T., & BIBBY J.M., (1979), *Multivariate analysis*, Academic Press Inc. (London) Ltd, New York, N.Y.

NETER J., WASSERMAN W & KUTNER M.H., (1983), *Applied Linear Regression Models,* Richard D. Irvin Inc Illinois.

NETER J. WASSERMAN W & WHITMORE G.A., (1982), *Applied Statistics*. Second Edition, Allyn and Bacon, Inc, Boston, Massachusetts

OVERALL J.E. & KLETT C.J., (1972), *Applied multivariate analysis*, McGraw - Hill Book Company, N.Y.

OWEN K. & Taljaard J., (1988), *Handleiding vir die gebruik van Sielkundige en Skolastiese Toetse van IPEN en die NIPN*, RGN, Pretoria

PENINSULA TECHNIKON, (1991), *Yearbook: School of Electrical- and Mechnical Engineering and Computer Data Processing*, Bellville

RAWLINGS J.O., (1988), *Applied regression analysis: A research tool*, Wadsworth & Brooks / Cole advanced books & software, Pacific Grove, California.

STOKER D.J. et al., (1985), *Investigation into Differential Entrance Requirements for Tertiary Educational Institutions*, An abridged version of Report WS-31, HSRC, Pretoria.

VAN ZYL T & MYBURGH C., (1991), *HSRC Catalogue Part 1: Psychological and Proficiency Tests: Individual and Group Intelligence Tests, Personality Tests and Aptitude and Proficiency Tests*, HSRC, Pretoria.

WEISBERG S., (1980), *Applied Linear Regression*, J Wiley & Sons, New York

WETHERILL G.B. et.al., (1986), *Regression analysis with aplications*, Chapman & Hall , London / New York.