An Analysis of Differential Aptitude Tests as a Predictor for Learner Performance in First Year Biosciences.

By
Jeffrey Franklin Finnie
812818031

Submitted in Partial Fulfilment
For the Degree of:
Master of Education (Higher Education)

In the
School of Education, Training and Development

University of KwaZulu – Natal
Pietermaritzburg

December 2004

Preface:

The experimental work described in this dissertation was carried out in the School of Botany and Zoology, University of KwaZulu-Natal, Pietermaritzburg, from February 2004 to December 2004, under the supervision of Kathy Luckett of the School of Education, Training and Development. University of KwaZulu-Natal, Pietermaritzburg.

The studies have not been submitted in any other form to another University and, except where the work of others is acknowledged in the text, are the results of my own investigation.

Jeffrey Franklin Finnie
January 2005

I certify that the above statement is correct

K M Luckett
(Supervisor)

Foreword

"SAUVCA has initiated a future-oriented project that will see higher education institutions pool expertise to develop a common set of benchmark tests. Institutions have reached a landmark decision to develop academic literacy, numeracy and mathematical tests that will tap these critical competencies needed for success in higher education studies. Over a period of three years benchmarks will be set on the tests that clearly spell out the standard required of entering students. By the end of 2008—at the same time as the Further Education and Training Certificate (FETC) will replace the current Senior Certificate school-leaving qualification—higher education institutions will include in their Admissions Policy how test results and benchmarks will be used to guide decision-making with regard to admissions, placement and academic support. This initiative speaks directly to the goals of equity and efficiency expressed in the National Plan for Higher Education, and the sector's concern with system linkages between the FET and HE sectors. There are three major advantages to setting benchmarks or entry-level standards for higher education study:

• Applicants, parents, schools, colleges and the public at large will be very clear on the expected entry levels to higher education institutions and how they are assessed;

• Higher education institutions will be clear on the entry level competencies of their students and what needs to be entailed in "responsive" curriculum programmes; and

• The formal schooling system will receive useful feedback on the critical competencies students need to develop in order successfully to engage with higher education study.

Most higher education institutions have over the past decade and more administered tests to first-year applicants to evaluate levels of proficiency. Test results are typically used in conjunction with Senior Certificate results to guide admissions decision-making, programme placement, and recommendations regarding academic support. However, the reality of differing entry tests has made it difficult if not impossible for the sector to provide consistent and systematic feedback to its target constituencies (particularly the school sector) about the extent to which learning experiences are adequately preparing learners for higher education studies.

It is clear that national benchmarks can play a critical role in higher education admission decision-making.

At the same time it is important that benchmark tests do not undermine the value of the current and future school-leaving qualifications. For example, Senior Certificate results provide useful information on the knowledge base and achievement levels of entering students in various subjects, and so will the future FETC. However, the value of the new schools curriculum and FETC in preparing students for higher education studies will be unknown until the first cohort of students entering with an FETC completes their higher education studies. At that stage the ability of the benchmark tests to predict academic success would have been established".

Press Release. Released by: SA Universities Vice-Chancellors Association
4 December 2004

Abstract:

This research project aims to establish the predictive validity of a set of aptitude tests for the first year Biosciences programme at a merging South African University. The study aims to address the problem of selection and placement to higher education and also to suggest how the results of aptitude tests might be used to inform curriculum development at first year level.

The Differential Aptitude Test (DAT-L) was conducted on the 2004 Biosciences student cohort (228 students). The instructions that were given to the students followed the instructions laid out in Owen and Vosloo (1999). Normally there are ten tests in this battery of tests, however only eight of the tests were given to the students. The tests included; Vocabulary, Verbal reasoning, Non-Verbal reasoning: Figures, Calculations, Reading Comprehension, Comparisons, Price Controlling and Memory.

Exam results of the participating students were collected. These included; Final Biosciences101 mark, Class mark (Biosciences101), Practical exam mark (Biosciences101) and Theory exam mark (Biosciences101). The final marks for Physics (mixed modules), Mathematics (mixed modules) Computer Sciences (CSCI 103) and Chemistry (CHEM131) were also collected. The indicator for the concept of academic potential in the Biosciences was taken as the marks achieved during the course of the first semester.

Biographical information from the Student Management System (SMS) was also recorded this included; Race, Gender, Home language and Matric points.

These various categories were compared against the students' performance in DAT- L tests.

The average for the stanine points for a student of university "quality' is seven, according to Owen *et al.* (2000), while the average of this cohort of students is 5.16. The correlation for the different DAT tests when compared with the final result for Biosciences shows that the best correlation is with test 1 at r = .47292 (Vocabulary) while the worst correlation is with test 6 at r=.24722 (Comparison). The sequence of correlation is from Vocabulary through Reading Comprehension, Verbal Reasoning, Calculation, Memory, Price Controlling, Non-Verbal Reasoning: Figures, and finally Comparisons.

The correlation of the average of all tests and the final result is a reasonable r = .50396. The best correlation with DAT and other subjects is Computer Science with r = .41165, with a declining correlation between Physics and the mark of r = .34085 followed by extremely poor correlations for Chemistry and Mathematics of r = .20313 and r = .08700 respectively. The best correlation that was obtained during this research was with the correlation between matric points and the average of all the DAT tests. While the correlation with DAT and the Final mark for Biosciences is r = .50396 that for the matric points is r = .57150.

Females attained a significantly higher average on the DAT tests but the difference between female and male on their final results in Biosciences is not significant. White students perform significantly better than the other race designations when it comes to the final result. Within the categories African, Indian and Coloured there is no significant difference. When it comes to the different DAT tests, Africans perform significantly worse than the other three groupings and the Indian grouping did significantly worse when compared to Whites.

If we were to use the DAT as a means of selecting students it would be fairly effective for biology but not for the other basic science subjects. The fact that the test would have to be administered to all applicants applying to do first year in the Science and Agriculture Faculty makes the choice of this set of tests questionable. Should the Faculty require a means of determining selection into biological subjects then the tests show great promise.

The pursuit of greater fairness, validity and reliability in selection is an ongoing quest.

Contents

Acknowledgements

I would like to thank Kathy Luckett for supervising this dissertation. Thanks are due to Prof John Akhurst for the original idea and for help, enthusiasm and support of the project. Prof Colleen Downs is gratefully acknowledged for assistance with the statistics and Carol Mitchell for providing the DAT-L booklets and scoring template.

Finally to Jeanette, Jocelyn and Ross, thank you for all the love, encouragement, support and patience.

1 Introduction

In selecting a research project it was important to me, that the project has relevance to my teaching. In addition any practical 'spin-off' from the research would be of additional benefit.

I have been teaching first year biology for the last fourteen years (and have been in charge of the Biosciences course for almost as many years) and have been fascinated with the performance of the learners in their first semester at university. Their ability has been a source of great delight and utter despair. In the past few years there has been a marked decline in the performance of the learners. The bell of "normal distribution" curve has slipped into the 50th percentile and the skills of the learners have also shown a marked decline. In my naivety, I though that it would be easy to design instruments to test the skills levels of the learners, however, after perusing the literature and speaking to a number of researchers in the field, I soon discovered that this is a misguided notion.

It is important to understand what academic skills learners "come with" on entering higher education. There is a great diversity (skills, background etc) in the learners that register for first year and it is important to take note of these differences and accommodate them as far as possible in ones' teaching. One way of determining what learners 'come with' is to use the diagnostic assessment test battery generally known as the Differential Aptitude Tests (DAT). The DAT (L) test was developed by Owen, Vosloo, Claassen and Coetzee (Owen *et al.*, 2000). Aptitude tests provide information on a learner's relatively strong and weak aptitudes and indicate also how her/his aptitudes compare with those of a comparable group - on the basis of which information certain inferences can be made. The DAT-L was constructed to provide learners with counselling and to give employers and educational institutions guidance, particularly with a view to determining potential and its (later) meaningful realization in the optimal utilization of educational and training facilities and the eventual development of high-level human resources. In choosing to test the validity of aptitude tests as my research question, I can, in part, build on the type of research with which I am familiar (quantitative research).

Brown (1995) states, that the approach to the study of science should display structural coherence, continuity and balance. As a component of a science curriculum, Biology makes an important contribution to the provision of breadth, balance and coherence (Brown 1995). Writing in a different context, Shayer and Adey (1981) have pointed out that because the sciences are different from one another, their teaching demands different approaches. They argue that while the methods of experimental design in biology are sophisticated in their demands, the concepts that are involved are capable of being understood at a number of different levels.

Activities in science classrooms should encourage the ability to plan and carry out investigations in which learners:

- Ask questions, predict and hypothesise;

- *Observe, measure and manipulate variables;*
- Interpret their results and evaluate scientific theory.

Through investigative activities in science learners are able to develop their ways of thinking about science and their understanding of its concepts and knowledge.

Can we predict what competences are required for successful study of the biological sciences and who would makes a good biologist? Probably not! However, it is important to attempt to establish mechanisms by which we can measure potential for the study of biology. The following are lists of skills that we attempt to teach in our first year biology classes.

Practical Skills

- Manipulative - manipulate simple apparatus; take accurate measurements and record them correctly
- Experimental - design experiments; formulate and test hypotheses; control variables; recognize experimental limitations and uncertainties; analyse and display data appropriately
- Drawing - make correct and careful observations; record observations by means of appropriate drawings (annotated); interpret observations

Thinking

- distinguish inferences from observations
- make appropriate approximations; estimate; and have a sense of orders of magnitude, interpret abstract representations e.g. graphs
- translate between biological phenomena and representations e.g. diagrams
- recognise limits of applicability e.g. of concepts
- identify assumptions underlying an argument and generate a logical argument

Problem-Solving

- tackle novel situations
- design a strategy
- execute a plan
- interpret a solution

Conceptual Understanding

- explain concepts in own words
- represent a concept in different ways
- apply correctly certain concepts in different circumstances

Reflection

- identify whether or not something is understood
- think about why a particular concept may be difficult to understand
- be aware of what sorts of factors help or hinder one's learning

Writing

- gather and synthesize information from a variety of sources
- write a short essay which integrates the operation (describe, discuss, evaluate etc.) and relevant content information

- describe an experiment, including the results and a summary of the main findings in the form of a practical report

Given that these are the skills we attempt to teach in first year Biology, it would be beneficial to develop a test that could measure who is best suited to studying Biology at university level.

Higher Education institutions are confronted daily with issues relating to the formulation of admission and selection requirements for students. If a method were found to predict academic success, this would not only minimise personal failure, but also bring about financial savings for both the student and the University (Swanepoel 2003). Selecting students for tertiary study is a worldwide problem.

Traditionally, scholastic achievements were used as admission criteria to gain access to South African tertiary institutions (Swanepoel 2003). However, in a changing South Africa, with its increasing emphasis on individual rights and its history of unequal educational provision, fair and equitable selection techniques for Higher Education are a priority (Swanepoel 2003). Placement tests will always be controversial. However, placement tests would be less controversial if they led to accurate predictions.

Differential aptitude tests (DATs) already exist; do they have predictive validity for first year biology in South Africa? Can these tests be used to predict successful academic study of Biology? Can the DAT tests be used to complement school leaving certificates in making judgments about selection and placement, particularly in light of the proposed Further Education and Training Certificate (FETC)? (see Foreword and Chapter 2). And finally can the results of the aptitude tests be used to improve teaching approaches and methods in first year Biology teaching?

2 Literature review

2.1 The problem

From 2008, the old-format Senior Certificate, with its grades, symbols, Matriculation points and Matriculation Endorsement requirements will vanish and be replaced with the Further Education and Training (FET) system. A new certificate, a Further Education and Training Certificate (FETC), replacing the Senior Certificate, will be awarded for the first time in 2008 as the South African school leaving certificate. The new FETC will be a full qualification (i.e. 120 credits) that will allow its holder access to higher education or entry into the workplace. South African Universities Vice Chancellors Association (SAUVCA) (2003) expresses concern that performance at the FETC exit-level is not expressed in terms useful to admission and selection to higher education institutions. 'The "criterion – referenced descriptors" are far too vague and general to be of use, as are the differences in levels or bands of achievement' (SAUVCA, 2003). SAUVCA further contends that the proposed system provides an insufficient fine-grained picture of learner performance upon which higher education can base its enrolment and placement practices. They suggest that this is a problem for higher education and other sectors of society as well. In addition Professor Schreiner, Acting Deputy Vice Chancellor (Academic) at the University of KwaZulu-Natal, suggests that the new system will adversely affect the selection process and the first year curriculum and a new calibration system to judge school results for admission of first-year students will have to be developed (Naude, 2003). It is further suggested that the minimum level of preparedness for higher education should be established and that universities should select the 'best risk' students (by developing our own entrance and placement tests) and by offering a range of differentiated access options (Luckett, 2004, pers com).

In the light of the changes to be introduced when assessing and grading school leavers' performance, research into the validity and reliability of aptitude tests and comparing these with first year Biosciences learners' performance may offer an alternative or additional means of assessing learners' potential for selection and placement in the Biosciences.

In the early 90's research was carried out at the Universities of Cape Town and Stellenbosch, where the abilities of first year science students were compared to their performance. This research was conducted at institutions that were highly selective in their student intake. The research is also no longer relevant in light of the changes in student intake that have taken place in the last few years (Jordaan 1990). The proposed research will attempt to produce similar research but make it relevant to the current situation in higher education, bearing in mind that the first year biology students at UKZN are not the result of a highly selected intake.

The University of Cape Town has since developed an 'alternative admissions' test. These tests are used to complement matric points and can be used as entrance criteria for disadvantaged learners.

Lowe & Cook (2003) explain that in a pilot survey of science students conducted by Cook and Leckey student study habits formed in secondary school persist to the end of the first semester of university life. Such a conclusion indicates that students are not bridging the gap between school and university quickly and effectively (Lowe & Cook, 2003).

In the light of this, this research project also attempts to address the question: can the results of Differential Aptitude Tests assist in curriculum development in first year Biosciences courses and in the making of appropriate pedagogical responses to students learning needs associated with bridging the gap between school and university in the context of a first year Biosciences course?

2.2 Aptitude tests

"Aptitude is the potential a person has which will enable her /him to achieve a certain level of ability with a given amount of training and /or practice" (Owen *et al.*, 2000).

The purpose of an aptitude test is to measure an individual's potential for obtaining a certain goal. In this case the goal is the successful completion of a 1st year biology course at university. It is further held that aptitude, together with interest, attitude, motivation and other personality factors, will to a large extent determine the ultimate academic success of the student. "It is assumed that a learner's ability to reason abstractly on the basis of verbal and non-verbal material, her /his memory for meaningful material, her /his quick and accurate visual and spatial perceptual abilities, and her /his ability to see and apply arithmetical, mechanical and physical principles, will give an indication of the degree to which he /she will be able to complete an academic, technical or practical course successfully" (Owen *et al.*, 2000). It is important to remember that a biologist must be capable of integrating all these facets in their career.

Super (1949) reported that Spearman speculated on the fact that there is only one intelligence factor "g" (IQ), a number of group factors such as word fluency and goodness of character and many specific factors, which are found in one test or situation (Super, 1949). Spearman further highlighted other special aptitudes such as: number, visualization, memory, word fluency, verbal relations, perceptual speed and induction. These special aptitudes form part of most aptitude tests that are in use these days.

Aptitude tests are designed to determine the potential a person has for the learning of a specific vocation or the ability to acquire specific skills (Hull, 1926). As a result aptitude tests are mainly used for educational and vocational guidance (Nisbet, 2000). Scores obtained using aptitude tests provide many pieces of useful information. Thus aptitude tests give information about a students strong and weak aptitudes and this gives an indication as to how their aptitudes would compare with those of a comparable group (Owen *et al..*, 2000).

Aptitude tests have the ability to measure a student's performance across a wide range of mental capabilities, similar to that of intelligence tests. However, aptitude tests also allow for the testing of more specialized capabilities such as verbal and numerical skills that can be used as a predictor for scholastic performance. Irrespective of the curriculum or background a student has experienced, aptitude tests can inform us as to what the student brings to the task in hand (Macklem, 1990, Owen *et al.*, 2000).

When completing an aptitude test, the person being tested is introduced to new and different situations in which he / she cannot rely heavily on previous training, but is dependent on his / her own ability to come up with a solution to the given problem. As a result of this feature of novel problem solving, aptitude tests are useful for the prediction of future achievement (DeVilliers, 1977). DeVilliers (1977) contends that it is nearly impossible to measure aptitude directly as these tests measure an individual's abilities at any one given moment.

Group aptitude tests have their use in three major areas according to Macklem (1990) these are:

- Instructional: Test results can be used to adapt the curriculum to match the level of the students or can be used to design tasks for students with different abilities. Test results can give the teacher realistic expectations of the capacity of the student.
- Administrative: Test results can be used to identify students into groups that require extra attention or those that need additional stimulation. The results can also be used in predicting job training performance.
- Guidance: Test results give a realistic expectation of academic performance.

Nisbet (2000) echoes the above, and in addition believes aptitude tests: are good predictors of future scholastic achievement; provide ways of comparing performances with others in the same or similar situation; provide a profile of strengths and weaknesses; assess differences among individuals; and can be used to uncover hidden talents.

A major advantage of group aptitude tests such as DAT-L, is that it can be given quickly and inexpensively to large numbers of students.

In considering an aptitude test, it is never a simple matter to declare it valid or not. It is also necessary to look at the strength of its predictive validity, and consider whether its predictive ability can be generalized across a range of situations. This is the premise on which the proposed research is based. To be more specific, the results of the aptitude tests will be compared with those of first year biology assessments to determine the extent to which the scores obtained in the DAT tests can be related to the successful completion of Biosciences 101.

It has been reported that in the United States the SAT can be useful in predicting college attainment, although others have questioned this. In the light of concern expressed that performance at the FETC exit level will not be expressed in terms useful to admission and selection to higher education institutions, can aptitude tests that have been specifically adapted for South Africa, be used with predictive validity in a local context?

2.3 Use of SAT to predict success in the USA

The following discussion is heavily based on the work of (McDonald *et al.* (2000), in which they report the findings of a literature review on aptitude testing for university entrance. The work was commissioned by The Sutton Trust and conducted by the National Foundation for Educational Research (NFER) in the United Kingdom.

In 1900, in an attempt to reconcile the considerable variation in entrance procedures between colleges in the United States, the College Board was formed. The test that is recognised as the SAT was first offered in 1926. 'SAT' originally stood for Scholastic Aptitude Test, but as part of a revision in 1994 it was renamed the Scholastic Assessment Test [Including the word 'aptitude' in the test title implied that the SAT measured innate differences]. Commenting on the lower scores obtained by Black students, Jencks (1998) observed that this 'suggested that blacks might suffer from some sort of innate disability whereas using the word 'assessment' removed the innate connotation]. Although the SAT has been relabelled as an 'assessment' test rather than a test of 'aptitude', it is still generally conceived as a test of academic aptitude

The SAT is not the only selection instrument for higher education in the United States, as the American College Testing Program has offered the American College Tests (ACT) since 1959. Whilst SAT was developed more as an aptitude test, the ACT measures the knowledge, understanding, and skills that have been acquired (American College Testing Program, 2000).

A second series of tests (SAT II: Subject Tests) has also been introduced. These have been developed from the College Board's Achievement Tests, and cover the major areas of academic study: English, math, social studies, natural sciences, and languages (McDonald *et al.*, 2000).

The SAT gives two scores: verbal and math. Using these scaled scores it is possible to obtain an indication of how the verbal or math scores for an individual student compare with the scores of others (McDonald *et al.*, 2000).

Current verbal SAT consists of:

- analogies measuring reasoning skills and knowledge of vocabulary

- sentence completions measuring logical relationships and vocabulary in the context of the sentence
- four reading passages of between 400 and 800 words each, with questions assessing extended reasoning skills, literal comprehension and vocabulary in context. Two of the passages are paired and complement each other in some way (e.g. represent differing sides of an argument), and so support questions relating to differing styles of writing or points of view, and one passage is described as being 'minority-relevant'.

Current math SAT consists of:

- quantitative comparisons and open-ended questions, or 'student-produced' responses.

Overall, in 1999 it was reported that about 85 per cent of colleges generally required students to take admissions tests (Schneider and Dorans, 1999). Schaffner (1985) reported that 55 per cent of colleges routinely considered admissions tests when reaching decisions, whereas 21 per cent did not require them or rarely used them. Thirteen per cent said they were required or recommended but rarely actually used in admissions. Sixty-five per cent of institutions reported high school achievement as being 'very important' or 'the single most important factor' in reaching decisions, whereas only 42 per cent thought this of entrance test results.

Do aptitude tests tell us any more than we already know? Scores from tests such as the SAT have been generally observed to predict achievement in college to about the same extent, or slightly less than other information on students such as high school grades and class rank (McDonald *et al.*, 2000). In actually using SAT or ACT results, the most common approach involves combining these with information such as high school class rank (HSCR) and grade point average (GPA). According to Crouse and Trusheim, (1988), high school information has consistently been identified as the best predictor of academic progress in college. Bridgeman et al. (2000) reason that SAT is useful in addition to high school record. Overall, predictive statistics show that adding the SAT to the high school record improves a college's estimates of students' academic performance (McDonald *et al.*, 2000).

In order for test scores to be used in guidance, it is necessary that they have some predictive validity. Stricker *et al.* (1996) reported a median validity of 0.42 for the prediction of first-year grades indicating its usefulness for guidance. Table 2.1 correlates first-year grades with SAT and HSGPA, in the USA.

McDonald *et al.* (2000) report that, the association between the SAT and FGPA varied between colleges from a high of 0.72 to a low of 0.37. In Israel Beller (1994) found the PET to be a good predictor of FGPA, with the average correlations being 0.53 for liberal arts, 0.5 for sciences, 0.45 for social sciences, and 0.43 for engineering. In the United Kingdom, the mean

A-level grade was the best predictor of degree performance followed by school assessment with average correlations of 0.36 and 0.26 respectively.

As African American students are known to score lower on tests such as the SAT and ACT (Bridgeman *et al.*, 2000), the predictive validity of these tests has been questioned. Fleming and Garcia (1998) observe that predictions for Blacks' grades were far more variable than for Whites. They also noted that the reasons for this variation have not been adequately explained. Schmitt and Dorans (1990) found that Blacks and Hispanics were less likely to reach items at the end of the verbal section.

Table 2.1. Correlations of first-year GPA with SAT and HSGPA (Adapted from McDonald *et al.*, 2000)

	African American		Asian American		Hispanic/Latino		White	
	Males	Females	Males	Females	Males	Females	Males	Females
SAT verbal	0.23	0.29	0.24	0.26	0.19	0.29	0.28	0.30
SAT math	0.30	0.34	0.32	0.32	0.19	0.31	0.30	0.31
SAT total	0.34	0.37	0.36	0.37	0.34	0.37	0.33	0.35
HSGPA	0.34	0.29	0.28	0.26	0.30	0.29	0.38	0.34

Neisser *et al.*, (1996) observed that the lower scores of ethnic minority groups on IQ tests are not unique to America. In other countries which have disadvantaged social groups, similar differences have been observed (e.g. New Zealand, India). Across the world, children who are born into these minorities tend to do less well in education and drop out earlier.

Zeidner (1987) found that in people who took the test after they were 30, that test scores show least validity as predictors.

In studies over extended periods of time Morgan (1990), Fincher (1990) and Schurr *et al.* (1990) found that the predictive power of the SAT has declined over time. They argued that this may be due to colleges taking steps to reduce student failure such as providing greater assistance to students who need support.

Male and female high school graduates in 2000 had comparable scores on the verbal section of the SAT, but males performed better on the math. Comparisons between ethnic groups show that Whites outperformed all other groups on the verbal section of the SAT. This can be explained through the SAT being a test of developed reasoning abilities in English, the first language of most White test takers. The highest average score on the math section of the SAT was achieved by students in the ethnic group labelled Asians, Asian Americans and Pacific Islanders (McDonald *et al.*, 2000).

Bridgeman and Wendler (1991) reviewed gender differences in SAT and ACT math scores, and found that males tended to outperform females. However, this difference was not reflected in general college attainment, which tended to be equal for males and females, or where differences did occur, females showed higher attainment. Whites tended to outperform all other ethnic groups. The exception to this was in math where Asians, Asian Americans and Pacific Islanders had the highest scores on both the SAT and ACT. African Americans performed very poorly on both tests, having the lowest average scores of any ethnic group on both SAT sections and all of the ACT subtests (McDonald *et al.*, 2000).

Corley *et al.* (1991) found rural students had lower SAT verbal scores than urban students, but higher GPAs in high school and college. Students from more affluent backgrounds had higher SAT verbal and math scores, but little difference was seen in the high school or college grades of these groups. These authors conclude that as differences persisted through high school and college, they were not due to differing standards between rural and urban high schools.

Hyde and Linn (1988) conducted a meta-analysis of verbal ability on almost 1.5 million participants. From these studies, there emerged only a very small sex difference in verbal ability, with females tending to outscore males. In contrast to this trend, males tended to outperform females on the verbal section of the SAT. Hyde and Linn (1988) suggest that the more technical nature of the material in the verbal section favoured males. This meta-analysis also showed a decline in gender differences over recent years. Hyde and Linn argue this may have resulted from changing roles in society. The meta-analyses conducted provide evidence that the SAT is biased in favour of males. Harris and Carlton (1993) found that male and female students, who attained similar math scores, achieved these through different means. Males performed better on geometry and geometry/arithmetic items, whereas females performed better on arithmetic/algebra and miscellaneous items. They concluded that these findings support the view that males perform better on items that have a practical application, possibly because they see math as being more relevant to their daily lives than females.

Roznowski and Reith (1999) do not to support this idea. They found removing biased items to have little effect on predictive validity or overall measurement quality, and argued that too great a focus on removing biased items can be 'unnecessarily limiting to test development'. Neisser *et al.* (1996) also noted that the quality of the school is an important factor, as general skills such as problem-solving, the ability to concentrate and motivation can all be passed on through schools and have important effects on subsequent learning. Socio-economic status is therefore likely to be a contributory factor to score differences, particularly as certain ethnic minorities may be over-represented in the lower social classes in America and many other western countries.

Rosser (1989) contrasts females' consistently lower scores on the SAT with data from the College Board's validity studies, which showed female first-year college attainment to be as good as or better than that of males showing that the SAT does not adequately predict first-year college performance (Rosser, 1989).

High school record and admissions test scores are able to predict university grades to a modest degree in the USA. There has been considerable controversy over the fairness of tests such as the SAT, particularly in relation to sex and ethnic score differences. The SAT acts as a measure against which high school performance can be judged. The effect of removing the SAT could therefore be to lower standards (Myers, 1997). Overall, predictive statistics showed that adding the SAT to the high school record improved a college's estimates of students' academic performance, with predicted college grades being a fair approximation of actual attainment. They are less successful in predicting retention, as failure to complete a course is often not determined by academic reasons (Myers, 1997).

Braun *et al.* (1986) and Ragosta *et al.* (1991) found for those with learning difficulties and physical and visual impairments, a combination of the SAT and GPA was judged to be a good predictor of college performance. Students with hearing impairments performed poorly on the SAT, and when combined with GPA this tended to under-predict the performance of these students. Prediction was generally seen to be not as good for disabled compared to non-disabled students.

Wilmouth (1991) has criticized the format of the SAT. Wilmouth contends that students may be rewarded for only partial knowledge; this may be as a result of guessing or elimination that can be used as a mechanism to answer the majority of questions, because the format of the SAT is multiple-choice.

What would happen if the SAT were to fall away? McDonald *et al.* (2000) have summarized the following points.

- similar GPAs from very different schools may not mean that same thing,
- without results from a standardized test such as the SAT, decisions may be made on the basis of the prestige of the high school attended and judgments as to whether a student is likely to 'fit in' to the college,
- the increased emphasis on high school grades may mean that students opt for easier courses to boost their GPA, and their tutors may adopt more lenient grading procedures, leading to rampant grade inflation.

Wilmouth (1991) argues that the SAT effectively provides a check on high school grading, so preventing this becoming more lenient.

2.4 Aptitude tests in Sweden and Israel.

Recently aptitude tests were introduced in Sweden and Israel as a result of the increasing number of students applying to study at universities. This has meant that selection into higher education has been necessary. Although the tests used in these countries may not provide a perfect solution to this problem, they do provide a defensible answer to the problem of selection (McDonald *et al.*, 2000). In Israel the test that was developed is called the Psychometric Entrance Test (PET). PET measures various cognitive and scholastic abilities to estimate future success in academic studies (Beller, 1994). PET assesses two factors, which correlate with SAT verbal and math. PET scores generally predict first-year degree performance better than high school grades. (McDonald *et al.*, 2000; Beller, 1994).

In Sweden the tests are known as SweSAT, a number of similarities can be drawn between the SweSAT and the SAT. Males tend to score higher on the SweSAT than females, as they do on the SAT (McDonald *et al.*, 2000).

2.5 Aptitude tests in the UK.

In the United Kingdom in the sixties and seventies a test known as the Test of Academic Aptitude (TAA) was developed. This added very little to the prediction after A-levels had also been entered with school assessment and O-levels. Sex differences in prediction were noted, particularly for science courses. Overall, a mean A-level grade was the best predictor of degree performance followed by school assessment (McDonald *et al.*, 2000).

Chapman (1996) explores the link between A-levels and degree results in eight subjects over 21 years. The strongest links were seen for biology (0.47) and weakest for politics (0.23). In the fifties Williams (1950) investigated the association between subject marks and first-year university performance. Prediction in science subjects was seen to be better than in arts. High positive correlations were seen between some subjects at matriculation and first-year degree attainment (Biology (0.77)), but many commonly studied subjects were predicted poorly from matriculation results (English (0.33), Physics (0.33)). Mellanby *et al.* (2000) shows that verbal reasoning test scores are a moderate predictor of degree class.

A generally applicable method of assessing aptitude may be desirable. Tests which have been identified as suitable for this purpose are based on the science reasoning tests developed by Shayer and Adey (1981). Validity work has been conducted on these tests, and they are known to be a good indicator of general intellectual functioning and to predict academic attainment.

Admissions tests would be more desirable if they could assess characteristics important for the successful completion of degrees, which are not measured by A-levels. McDonald *et al.* (2000) conclude that for the United Kingdom it is fundamental that further research is conducted into predictive validity, before any introduction of a national university admissions test.

Research has suggested that personality factors may be useful in predicting college attainment, but far more work is needed before there is sufficient evidence for these to be used in the admissions process.

2.6 The South African Context

According to Behr (1985) and Phala (1992), most tertiary institutions in South Africa have introduced an admission rating system, known as the Swedish formula, in which the symbols obtained in the matric examination, are weighted. The derivation of the formula for the admission to the University of KwaZulu-Natal is unknown (see Appendix 2). Given the current problems in education and the environmental constraints of the majority of applicants, the Swedish formula can no longer be used as the sole selection instrument (Swanepoel, 2003).

School performance refers to a pupil's performance in the different exams during his/her high school career with emphasis on matric or Senior Certificate exam. This performance is used as the legal criterion for University entrance (Swanepoel, 2003). Behr (1985), Bokhorst *et al.* (1992) and Van Wyk and Crawford (1984) have reported that successful first-year students obtained significantly better matric symbols than did unsuccessful students. Botha and Cilliers (1999), contend that although matric symbols are a poor predictor of academic success, these are the most often used criteria to determine entrance to tertiary institutions. Mitchell and Fridjhon (1987), as well as Botha and Cilliers (1999) mention that the matric average is still the most conclusive criterion. However, Miller (1992) feels strongly that matric results cannot be used as the sole predictor for entrance to tertiary institutions.

Kotecha (2001) contends that the Senior Certificate fulfils none of its three key functions adequately:

- it is not a reliable predictor of competencies required in the work place;
- it is not a reliable predictor of competencies required in HE or a predictor of probable success in HE; and
- it does not serve the needs of fully half of all learners attempting the exam.

Kotecha (2001) continues to argue that although the Senior Certificate should enable universities to screen and select their potential students, it is considered regressive and has a disproportionate impact on the form and content of learning and teaching.

Huysamen (1996; 1999; 2000) and Mitchell and Fridjhon (1987) feel strongly that it would be unfair to apply matric results or use the results of aptitude tests for those applicants who had inferior schooling as well as for those of who had privileged schooling.

Tertiary education demands acquired skills from applicants (Barnard, 1992 and Kilfoil, 1999). Botha and Cilliers (1999) state that, scholastic performance usually provides little information on the intellectual potential and aptitudes, especially for black applicants and their ability to succeed in tertiary studies. Kilfoil (1999) and, Botha and Cilliers (1999) confirm the unreliability of school results and stress that assessment of tertiary applicants is becoming increasingly important.

The National Commission on Higher Education (NCHE) in 1996 recommends that additional selection criteria in addition to the Further Education Certificate (FEC) be implemented (Swanepoel 2003).

Griesel (2001) reports, that institutions use a range of strategies to increase participation, in addition to school-leaving performance results, these can be grouped into three broad categories:

- once off testing that typically includes an assessment of English and/or Mathematics competence or proficiency; and, in other cases, batteries of psychometric tests;
- assessment broadly within a developmental or interactive approach that includes interviews, portfolios, biographies, and assessment strategies that mirror educational intervention; and
- school/community outreach programmes.

Although the ideal is to increase access and admission to Universities, Griesel (2001) has listed a number of constraints that prevent increasing participation from all sectors of the population.

- Large numbers of learners who are under prepared;
- Concerns about the sustainability of interventions;
- And related, the financial support required by students and needed for interventions;
- Institutional capacity in terms of physical and human resources.

2.7 Differential aptitude test (L) [DAT (L)]

The DAT (L) test was developed by Owen, Vosloo, Claassen and Coetzee. The following discussion is primarily a synthesis of relevant information from the manual for the DAT–L tests (Owen *et al.*, 2000). Aptitude tests provide information on a learner's relatively strong and weak aptitudes and indicate also how her/his aptitudes compare with those of a comparable group - on the basis of which information certain inferences can be made. On this basis deductions and predictions can be made about possible success in specific fields of study or training (Owen *et al.*, 2000). It is assumed that aptitude, in conjunction with interests, motivation and other personality characteristics to a large extent determine the ultimate educational/career success of a person.

With an aptitude test we want to establish whether a person has the ability to carry out a task in the future if he/she in the meantime receives the necessary training. In other words we want to establish whether the person has the necessary learning potential in a specific field to enable her/him to achieve success after training (Owen *et al.*, 2000).

Although aptitude tests in general are good predictors of school achievement, the correlations between aptitude tests and career success and satisfaction are not very high. One of the reasons for this is that preparation for entry into an occupation and later achievement in that occupation often calls for slightly different abilities.

In interpreting the test results one should first ascertain whether the learner's general ability is such that he/she can benefit from further training. The single test that gives the best indication of the general intellectual level of the learner is Test 2 (Verbal Reasoning). However, a more reliable indication will be obtained by taking into consideration also the scores of Test 3 (Non-verbal Reasoning) (especially in the case of learners with a language backlog), Test 4 (Calculations) and Test 5 (Reading Comprehension). If the learner obtains a mean stanine of 7[1] or higher in these four tests together, he/she will fall in the top 23 percent of the population and can be regarded as university material; with a mean stanine of 3 or lower he/she will fall in the bottom 23 percent of the population and will probably find it difficult to progress beyond the school grade in which he/she currently is; with a mean stanine of 4, 5 or 6 he/she will fall between the 23rd and 78th percentile, in other words where the majority of her/his peer group find themselves - he/she can therefore be regarded as an average learner.

In order to compare the abilities of a learner in respect of the three broad categories, Academic, Commercial/Clerical and Technical, the tests of the DAT-L can be classified as follows:

<u>Academic</u>
Test 2: Test 3: Test 4: Test 5: Test 1: Test 8:
Verbal Reasoning, Non-verbal Reasoning, Calculations, Reading Comprehension, Vocabulary, Memory.
These tests indicate the general learning ability of the learner and are thus related to her/his performance at school.
<u>Commercial/Clerical</u>
Test 6: Test 7; Test 5: Test 4
Comparison, Price Controlling, Reading Comprehension, Calculations
<u>Technical</u>
Test 9: Test 8: Test 3: Test 4: (as recorded by (Owen *et al.*, 2000).
Mechanical Insight, Spatial Visualization *(3-D)* Non-verbal Reasoning Calculations (These tests were not used in this research)

Finer differentiation within a particular field or direction can be obtained through the use of individual tests or a combination of tests. For example, a comparison of the scores of Tests 5 (Reading Comprehension) and 4 (Calculations) will enable one to decide whether a learner is better equipped for occupations where language plays an important role such as in journalism or office correspondence, or for occupations where arithmetical work is important as in accountancy or bookkeeping. The tests do not make provision for factors such as interest, dedication, perseverance,

[1] Stanines are out of a possible nine, calculated with reference to norms tables.

temperament, personality, working circumstances, and so on, which can all have a major influence on job success (Owen et al., 2000).

A learner's performance in respect of abilities such as verbal reasoning, reading comprehension, memory concerning verbal material, and so on, cannot be properly measured if he/she has a language backlog.

The DAT-L battery of ten tests determines aptitude taking into consideration verbal, non-verbal, spatial, numerical, clerical, and mechanical and memory factors (Owen *et al.*, 2000). The DAT-L was designed to measure a limited number of basic abilities in order to make predictions about a wide variety of occupations (Owen et al., 2000). Only persons who have already received at least five years of school instruction in English or Afrikaans should be tested with the DAT Form L (Owen *et al.*, 2000).

The sample used to construct the norms tables is not an accurately reflection of the distribution of race in South Africa, and Owen *et al.* (2000) admit that there was not sufficient representation among black students across the educational spectrum. The black student samples were skewed as only rural 'self-governing' schools were used as participants and this sample is not representative. In fact, Owen *et al.* (2000) state that the norms furnished in the manual are subject to serious shortcomings because of the problems experienced during standardisation of the DAT-L these included unrest at many black schools and many matriculants were preparing for the final examination. The different tests in general, function best for White learners followed by Indian and Coloured (Owen et al., 2000).

Despite the above the DAT-L has acceptable content validity, construct validity and predictive validity for the correct sample (Owen *et al..*, 2000). The use of an inappropriate sample would lead to the test having no validity because of insufficient measure of aptitude (Owen *et al.*, 2000).

Language can easily become a source of test bias as language is regarded as the single most important moderator of test performance. It takes longer for students that have English as a second language to process information and as a result poor performance on the aptitude tests could be more a reflection of language difficulties than of weak individual ability.

Eight of the potential ten tests were used in this research, they are as follows:

Test 1: Vocabulary

The aim of the test is to measure the Verbal Comprehension factor, which can be defined as the knowledge of words and their meaning, as well as the application of this knowledge in spoken and written language.

The ability of a learner to recognise a word and to choose a synonymous word is regarded as a valid indication of her/his knowledge of the meaning of words and as a valid criterion for the Verbal Comprehension factor.

Test 2: Verbal Reasoning

The aim of the test is to measure an aspect of General Reasoning (R) on the basis of verbal material. The test rests on the assumption that the ability to determine relationships, to complete word analogies, to solve general problems requiring logical thought, as well as a person's vocabulary background, is a valid indication of an aspect of General Reasoning.

Test 3: Non-verbal Reasoning: Figures

The aim of the test is to measure an aspect of General Reasoning (R) on the basis of nonverbal material. The test rests on the assumption that the ability to see the relationships between figures and, by analogy, to identify an appropriate missing figure, as well as, following the changes that the figures of a figure series undergo, to deduce the work principle and to apply it again, is a valid indication of an aspect of non-verbal reasoning ability.

Test 4: Calculations

The aim of the test is to measure the arithmetical ability of the learner. The test rests on the assumption that the learner's ability to do mechanical calculations and to solve arithmetical problems with the help of the four basic arithmetical operations, namely adding, subtracting, dividing and multiplying, provides a valid indication of her/his arithmetical ability.

Test 5: Reading Comprehension

The aim of the test is to measure the learner's ability to comprehend what he /she is reading. In this test it is assumed that a learner's ability to choose the right answers to questions on prose passages a valid indication of reading comprehension.

Test 6: Comparison

The aim of the test is to measure Visual Perceptual Speed as a certain aspect of clerical ability, which consists mainly of the quick and accurate perception of differences and similarities between visual configurations. The test is based on the assumption that the ability to quickly and accurately indicate from five symbol groups the one that corresponds precisely with a given symbol group, is a valid indication of Visual Perceptual Speed.

Test 7: Price Controlling

This test is intended to measure a general aspect of clerical ability, namely the ability to look up data quickly and accurately. The test is based on the assumption that the ability to look up the prices of articles in a table quickly and accurately is a valid indication of success in numerous clerical tasks.

Test 8: Memory (Paragraph)

The aim of the test is to measure an aspect of the Memory factor (M) by using meaningful material. The test is based on the assumption that the ability to memorise meaningful material summarised in written paragraphs and then to correctly answer questions on the content of the paragraphs, is a valid criterion for measuring an aspect of Memory.

Only one type of deduced score or norm score has been calculated for DAT-L, namely stanines. The stanine or nine-point standard scale is a normalised scale with standard scores ranging from 1 to 9 with a mean of 5 and a standard deviation of 1,96

A stanine of 5 is for example obtained by the middle 20% of the population; this performance exceeds 40% of the population and is in turn exceeded by 40%. In the same way a stanine of 7 is obtained by 12% of the total population; it exceeds 77% and is exceeded by 11 %. The stanine scale is ideal for comparing different learners' scores, and for comparing the different scores of the same learner. Stanines are more stable than other types of standard scores and percentile ranks because they usually represent broad intervals of unprocessed scores. Furthermore there is less likelihood of misinterpretations and overgeneralizations of small and unreliable differences between unprocessed scores, except in cases where the scores lie near the boundaries of the different stanines (Owen *et al.*, 2000).

Table 2.2 Description of the stanine scale (Owen *et al.*, 2000).

Percentage tests	Stanine	Description
Lowest 4 %	1	Very poor
Next 7%	2	
Next 12%	3	Poor
Next 17%	4	
Middle 20%	5	Average
Next 17%	6	
Next 12%	7	
Next 7%	8	Good
Highest 4 %	9	Very good

It is important to note that the norms furnished in this manual are subject to serious shortcomings. Because of the problems experienced during the standardization of the DAT-L, the norms should be regarded only as a broad guideline and rough indication of possible inter-individual and intra-individual differences.

2.8 Conclusion

In the South African context, Senior Certificate results are increasingly less valid and reliable for subjects taken on the standard grade (SG), but performance on subjects taken on the higher grade (HG) is a reasonably accurate predictor of academic potential. Furthermore, research has shown that Senior Certificate results are better predictors of academic progress for students from educationally advantaged backgrounds than for those from educationally disadvantaged backgrounds (Foxcroft, 2004).

When higher education institutions use the FETC results as selection criteria for the first time with the 2009-intake of students, unsuccessful applicants could sue them and they may have no empirical data to support their rejection of an applicant with an FETC (Foxcroft, 2004).

The higher education sector urgently needs to benchmark the FETC against the critical competencies required of students for entrance to higher education programmes. This can be achieved by developing benchmark tests to test for academic literacy, numeracy and mathematical competencies that are critical for success in Higher Education studies. By using the results of the benchmarking tests, Higher Education institutions would become more definite about their entry requirements (Foxcroft, 2004).

However, the Employment Equity Act states that the use of tests is prohibited unless they have been scientifically shown to be valid and reliable. The International Test Commission (ITC) goes further and states that tests that have inadequate or unclear supporting technical documentation should be avoided (Foxcroft, 2004).

Differential aptitude tests (DATs) already exist; this study seeks to discover the extent to which they have predictive validity for first year biology students in South Africa? Can these tests be used to predict successful academic study of Biology in South Africa? Can the DAT tests be used to complement school leaving certificates and to assist South African higher education institutions in making judgments about selection and placement?

3 Methodology and Research Method

3.1 Methodology

Orlikowski and Baroudi (1991), following Chua (1986), suggest three research paradigms, based on the underlying research epistemology: positivist, interpretive and critical. This three-fold classification is the one that is to be adopted in this piece of research. However it needs to be said that, while these three research paradigms are philosophically distinct (as ideal types), in the practice of social science research these distinctions are not always so clear-cut. There is considerable disagreement as to whether these research "paradigms" or underlying epistemologies are mutually exclusive opposed or can be accommodated within the one study.

Due to the nature of this research, the approach to the interpretation of the results was primarily quantitative, thus the study is undertaken from within the positivist paradigm. A positivist ontology generally assumes that reality is objectively given, whilst a positivist epistemology assumes that reality can be described by measurable properties which are independent of the observer (researcher) and his or her instruments. Positivist studies generally attempt to test theory, in an attempt to increase the predictive understanding of phenomena, as is the case for the research question posed in this study. Orlikowski and Baroudi (1991) classify research as positivist if there was evidence of formal propositions, quantifiable measures of variables, hypothesis testing, and the drawing of inferences about a phenomenon from the sample to a stated population, as proposed in the research questions above.

Interpretive and Critical researchers assume that social reality is historically constituted and that it is produced and reproduced by people. Critical researchers recognize that their research is constrained by various forms of social, cultural and political domination (Mouton, 1996). The main task of critical research is seen as being one of social critique, whereby the restrictive and alienating conditions of the status quo are brought to light. It is noted that there has been considerable controversy over the fairness of tests such as the Scholastic Assessment Test (SAT) in the USA particularly in relation to gender and ethnic score differences. For example evidence suggests that both the verbal and math sections of the SAT are somewhat biased in favour of males, whilst groups such as African Americans score about one standard deviation lower on the SAT than Whites. Furthermore, the SAT under-predicts female attainment. These factors were analyzed as part of this research project, thus including a critical dimension in an otherwise positivist approach.

Today the competing paradigms of quantitative and qualitative research have become almost working partners in educational research. Many researchers today advocate a paradigm of choices that seeks methodological appropriateness as the primary criterion for judging methodological quality (Fierro, 2004).

Quantitative research methods are appropriate where quantifiable measures of variables of interest are possible, where hypotheses can be formulated and tested, and inferences drawn from samples to populations. Quantitative research is based primarily on positivist thought. Quantitative research is based on statements such as "anything that exists, exists in a certain quantity and can be measured." Custers (1996) quotes that "While Thorndike's statement from 1904 appears to be fairly innocent and direct, it staked an important philosophical position that has persisted in social science research throughout most of this century." (Custer, 1996). The ideals of quantitative research call for procedures that are public, that use precise definitions, that use objectivity-seeking methods for data collection and analysis, that are replicable so that findings can be confirmed or disconfirmed, and that are systematic and cumulative—all resulting in knowledge useful for explaining, predicting, and controlling, in the case of educational research, the effects of teaching on student outcomes (Fierro, 2004).

Quantitative methods of research are designed to be detached from, and independent of, a specific situation under study. In the design of such research, objectivity is a key criterion for internal and external validity. This allows the research findings to be applied to more than one population, which increases the generalizability of the research. Another reason for the preference of quantitative methods is that the conclusions obtained from such studies are considered to be more reliable and statistically valid (Mouton, 1996).

Studies aimed at quantifying relationships are of two types: descriptive and experimental. In a descriptive study, no attempt is made to change behaviour or conditions, one measures things as they are. Descriptive studies are also called observational, because one observes the subjects without otherwise intervening. In an experimental or quasi-experimental study, measurements are taken, then some sort of intervention occurs, then measurements are taken again (Hopkins, 2000).

The various research designs differ in the quality of evidence they provide for a cause-and-effect relationship between variables. Single cases and case series, which are of a quasi-experimental design, are the weakest. A well-designed cross-sectional or case-control study can provide good evidence for the absence of a relationship. But if such a study does reveal a relationship, it generally represents only suggestive evidence of a correlation. Confounding is a potential problem in descriptive studies that try to establish cause and effect relations. Confounding occurs when part or all of a significant association between two variables arises through both being causally associated with a third variable (Hopkins, 2000).

When the sample is not representative of the population, selection bias is a possibility. Failure to randomize subjects to control and treatment groups in experiments can also produce bias (Mouton, 1996).

To ensure statistical significance the sample size has to be big enough to ensure the detection of the smallest worthwhile effect or relationship between variables. The type of design chosen has a major impact on the sample size. Descriptive studies need hundreds of subjects to give acceptable confidence intervals (or to ensure statistical significance) for small effects (Hopkins, 2000; Mouton, 1996).

Positivist studies generally attempt to test theory, in an attempt to increase the predictive understanding of phenomena. The researcher must ensure that the measuring instrument is actually measuring the concept in question and not some other concept and that the concept is being measured accurately (Bailey, 1987).

The precision of measurement also has a major impact on sample size. Precision is expressed as validity and reliability. Validity represents how well a variable measures what it is supposed to. Reliability shows how reproducible measures are on a retest. Reliability of data implies the reliability of the ultimate source of the data. Accuracy of data implies the degree of 'correctness' of data based upon factors such as whether it is confirmed by data from a reliable source as well as the reliability of the original source of data (Hopkins, 2000; Mouton, 1996).

The study that I undertook was primarily of a descriptive design using quantitative methods, in which the students were compared with themselves. The analysis of the data focused primarily on correlations in order to assess the predictive validity of the DAT-L series of tests. The research did not involve any intervention, and so it cannot be considered a quasi-experimental design.

3.2 Research Method

The first question to be researched was: To what extent can the Differential Aptitude Tests be used as a predictor for learner performance in a first year Biosciences course? The data was be generated as follows: Aptitude Tests DAT-L -all Bioscience 101 ± 300 students were tested using the DAT test. The first year group was made up of learners directly from school as well as repeat students and students that have completed the Science Foundation Programme (SFP). The class composition was fairly evenly distributed with respect to gender and race. The resultant data was compared to the learner's final mark at the end of the semester. The aptitude tests were conducted during the first practical of the first semester and responses recorded on a computer score sheet. These were manually analyzed and assigned to various categories in the DAT tests. The various categories were compared against learners' marks in the exam and class marks (taken from the Student Management System (SMS)). This was further compared with their matric points to ascertain the extent to which the current matric points are a indicator of achievement in first year biology.

Academic potential in the Biosciences was measured in terms of the marks achieved during the course of the first semester. The DAT categories were compared with the learner's class mark (a result generated in a formative manner), categories were also related to the practical exam mark as well as the final theory exam. Results were statistically analyzed using STATISTICA ™. Comparisons were drawn between the various DAT categories and the respective test score and the performance of the learners in theory tests, practical tests, their performance measured in the formative assessment mark and their performance in the final summative exam.

As a result of evidence that aptitude tests show gender and race bias (see section 2.3), it is essential that the learners should be divided into racial and gender groupings in order to ascertain whether the same problems are encountered in the DAT-L tests. To ensure confidentiality of the results, no names or student numbers were divulged at any stage.

The aptitude test DAT-L was conducted during the first practical of the 2004 Biosciences student cohort. The context and content of the tests is not related to Bioscience and are considered decontexualised and are supposed to be 'neutral'. The instructions that were given to the students followed the instructions laid out in Owen and Vosloo (1999).

Normally there are ten tests in this battery of tests, however only eight of the tests were given to the students. As a result of the change in the number of tests the sequence and timing of the test as laid out by Owen and Vosloo (1999) was altered as follows:

Test 1	Vocabulary	10 minutes
Test 2	Verbal reasoning	15 minutes
Test 3	Non-Verbal reasoning: Figures	11 minutes
Break		10 minutes
Test 4	Calculations	20 minutes
Test 5	Reading Comprehension	15 minutes
Test 8	Memory learning time	6 minutes
Break		15 minutes
Test 6	Comparisons	8 minutes
Test 7	Price Controlling	8 minutes
Test 8	Memory	8 minutes

Responses recorded on a computer score sheet (Appendix 1).

Only persons who have already received at least five years of school instruction in English or Afrikaans should be tested with the DAT Form L, the language of testing was English[2].

Before the answer sheets were scored they were checked for two or more answers per question and all such answers were rubbed out. Where the student follows a certain fixed pattern[3] these sheets were discarded. Any incomplete answer sheets were also discarded. Two scoring stencils were used. The scoring process consists of placing the stencil correctly on the answer sheet and counting the pencil marks that can be seen through the

[2] Some of our foreign students have not had at least five years of school instruction in English.

[3] i.e. If all the A's were marked and nothing else, or the answers showed geometric patterning.

holes. Every mark that is visible through the scoring stencil counts for one mark. When all the tests had been scored the standard scores (stanines) are obtained by looking up the raw scores (i.e. the test scores) in the appropriate norm table. The norm table used was for grade 12 students.

Initially 262 students completed the test but 34 test sheets were disregarded as they displayed a fixed pattern, were incomplete or were incorrectly completed. This number although not overly large does represent a statistically significant number of samples.

Exam results of the participating students were collected. These included:

- Final Biosciences101 mark (Comprising; 34% Class mark, 33% Practical exam mark and 33% Theory exam mark)
- Class mark (Biosciences101)
- Practical exam mark (Biosciences101)
- Theory exam mark (Biosciences101)
- The final marks for Physics (mixed modules) Mathematics (mixed modules) Computer Sciences (CSCI 103), Chemistry (CHEM131) were also collected.

Biographical information from the Students Management Systems (SMS) was also recorded. This included:

- Race
- Gender
- Home language
- Matric points (not available for foreign students)
- Whether they completed the Science Foundation Programme (SFP)[4]
- Matric symbol for Biology

The first year group is made up of learners directly from school as well as repeat students and students that have completed the Science Foundation Programme. The class composition is relatively evenly distributed in respect to gender and race).

The various categories from the DAT-L were compared against students' marks in the exam and class marks (taken from SMS). The DAT-L categories analysed are outlined in the table below.

Table 3.1 Categories of tests and combination

Test No.	Type of test
Test 1	Vocabulary
Test 2	Verbal reasoning
Test 3	Non-Verbal reasoning : Figures
Test 4	Calculations
Test 5	Reading Comprehension

[4] The Science Foundation Programme (SFP) is a one year alternative access programme which provides an access route for underprepared students to enter the Faculty of Science and Agriculture. The programme is an outstanding example of the University of KwaZulu-Natal's strategic planning, initially to transform itself and to address the problems created by the educational inequities of secondary schooling in the apartheid era.

Test 6	Comparisons
Test 7	Price Controlling
Test 8	Memory
Test 2,3,4,5	Verbal reasoning, Non-Verbal reasoning :Figures, Calculations, Reading Comprehension
Test 1,2,3,4,5,6,7,8 (All the tests = Average)	Vocabulary, Verbal reasoning, Non-Verbal reasoning :Figures, Calculations, Reading comprehension Comparisons Price Controlling Memory
Test 1,2,3,4,5,8 (Academic)	Vocabulary, Verbal reasoning, Non-Verbal reasoning :Figures, Calculations, Reading Comprehension Memory
Test 4,5,6,7 (Commercial/ Clerical)	Calculations Reading Comprehension Comparisons Price Controlling

The various DAT-L categories were also compared with matric points to ascertain if the current matric points can be linked to the students' performance in the DAT tests.

Academic potential in the Biosciences was measured in terms of the marks achieved during the course of the first semester. The DAT categories compared with the learner's class mark (a result generated in a formative manner, the marks made up of class tests, an essay and eleven practicals). Categories were also related to the practical exam mark as well as the final theory exam.

Results were statistically analyzed using STATISTICA™ (and using analysis of variance; ANOVA RMANOVA) at the 95% confidence limits. Comparisons were drawn between the various categories and the respective test score and the performance of the learners during the 1st semester in theory tests, practical tests, the formative assessment mark and their performance in the final summative exam. For the other subjects the final exam mark was used for correlation purposes.

"Estimates of predictive validity are generally based on correlation coefficients or variations of these. Correlation coefficients indicate the extent of the association between two variables. A correlation of zero indicates that two variables are unrelated to each other and a correlation of one indicates a perfect, linear association between them. The stronger the correlation between an aptitude test and an outcome measure, the better its predictive validity is said to be" (MacDonald *et al.* 2000: 20). It is not sufficient to consider the simple association between aptitude test scores and outcome measures. Any statistics need to be presented in the context of other information.

As a result of evidence that aptitude tests show gender and race bias, the learners were divided into racial and gender groupings in order to ascertain whether similar problems are encountered in the DAT-L tests. This biographical data was also used to test against the DAT-L results.

4 Results

4.1 General.

Table 4.1 outlines the means of the various tests and variables that were recorded. Generally the marks for Bioscience have shown a consistent trend over the last few years. The most disturbing feature is the variation in the stanine points that are evident for tests one through to eight. The minimum mark score in all the tests is stanine one in all cases, a mark that a University student should not be getting. The maximum stanine points, however is to be expected. The average for the stanine points for a student of university "quality" is seven according to Owen *et al.* (2000), while the average of this cohort of students is 5.16. The average of the different combinations of the different tests (see Table 3.1) is below that of the average for all the tests. The other four subjects that were included for comparison purposes straddle the mark for Biosciences with Computer Sciences having the higher mark while the average for Physics is below 50%.

Table 4.1 The Means, Maxima, Minima and Standard Error for the Biosciences Class of 2004.

	N	Mean (%)	Minimum (%)	Maximum (%)	Standard Error
FINAL RESULT (BIOS 101)	228	61.71	28.00	91.00	0.65
CLASS MARK	228	60.10	39.70	87.80	0.62
THEORY EXAM MARK	228	57.39	26.80	92.80	0.86
PRAC EXAM MARK	227	68.21	36.00	92.50	0.63
TEST 1	228	6.82	1.00	9.00	0.16
TEST 2	228	5.93	1.00	9.00	0.14
TEST 3	228	3.48	1.00	9.00	0.13
TEST 4	228	5.03	1.00	9.00	0.13
TEST 5	228	3.93	1.00	9.00	0.15
TEST 6	228	5.73	1.00	9.00	0.15
TEST 7	228	5.21	1.00	9.00	0.15
TEST 8	228	5.17	1.00	9.00	0.14
AVE OF ALL TESTS	228	5.16	1.13	8.63	0.11
TESTS 2,3,4,5	228	4.59	1.00	8.25	0.11
TESTS 4,5,6,7	228	4.97	1.00	9.00	0.11
TESTS 1,2,3,4,5,8	228	5.06	1.17	8.50	0.11
CHEMISTRY FINAL MARK (CHEM131)	206	57.13	22.00	94.00	0.98
COMPUTER SCIENCES FINAL MARK (CSCI 103)	141	70.97	46.00	95.00	0.75
MATHS (MIXED MODULES)	155	64.74	29.00	100.00	1.19
PHYSICS (MIXED MODULES)	190	48.79	23.00	92.00	0.97

4.2 Correlation of the DAT-L to the final result.

The different DAT tests are correlated with the final result for Biosciences and graphically displayed in Figures 4.1 A-H. The best correlation is with test 1 at r = .47292 (Vocabulary) while the worst correlation is with test 6 at r=.24722 (Comparison). The sequence of correlation is from Vocabulary through Reading Comprehension, Verbal Reasoning, Calculation, Memory, Price Controlling, Non-Verbal Reasoning: Figures, and finally Comparisons. The sequence mostly makes sense except for the poor correlation with Non-Verbal Reasoning: Figures.

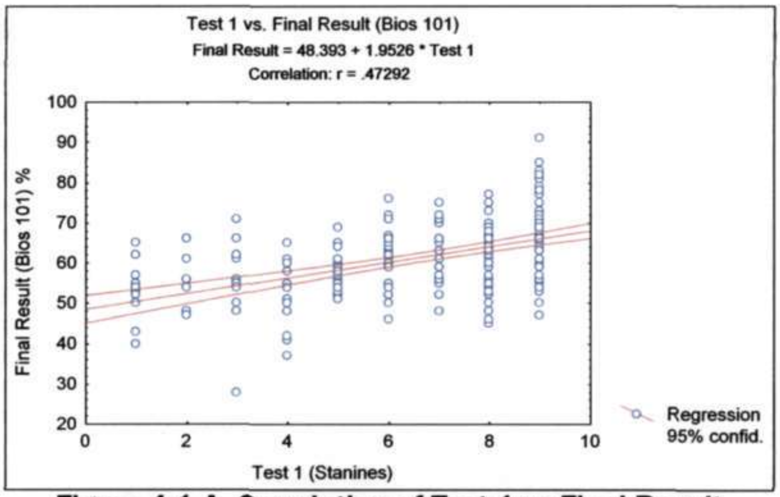


Figure 4.1 A. Correlation of Test 1 vs Final Result

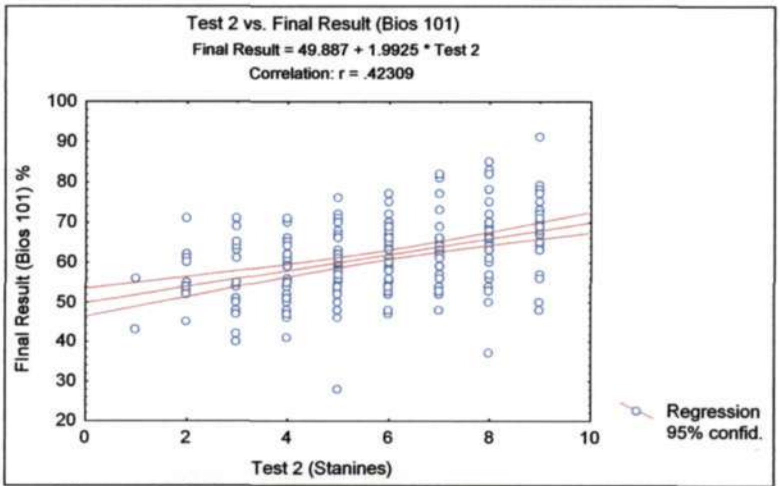


Figure 4.1 B. Correlation of Test 2 vs Final Result

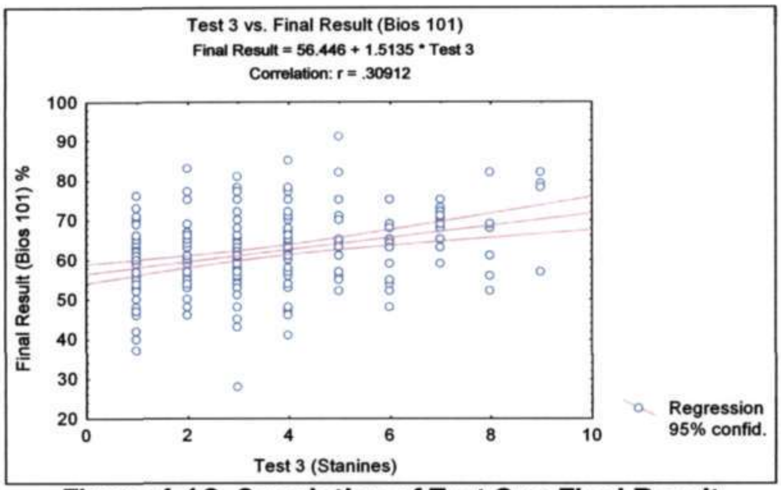


Figure 4.1C. Correlation of Test 3 vs Final Result

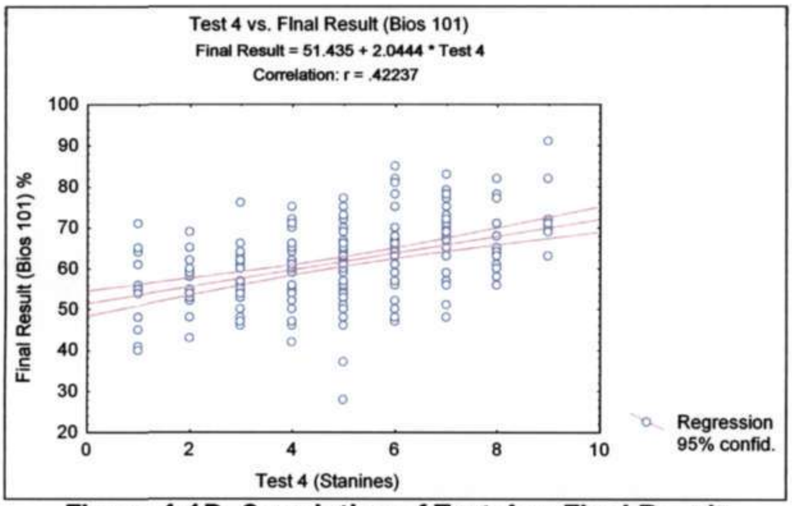


Figure 4.1D. Correlation of Test 4 vs Final Result

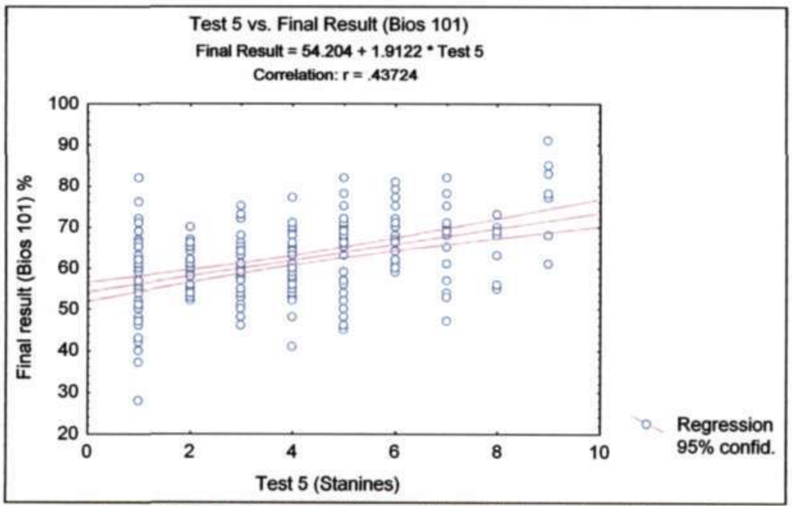


Figure 4.1 E. Correlation of Test 5 vs Final Result

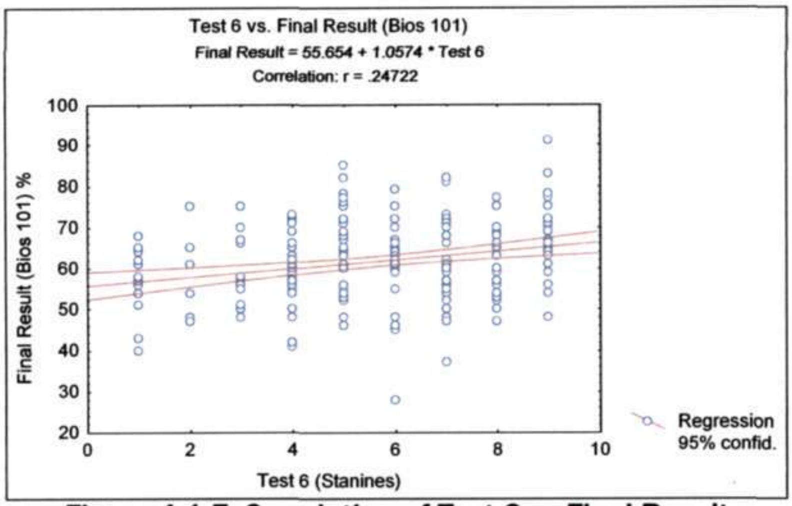


Figure 4.1 F. Correlation of Test 6 vs Final Result

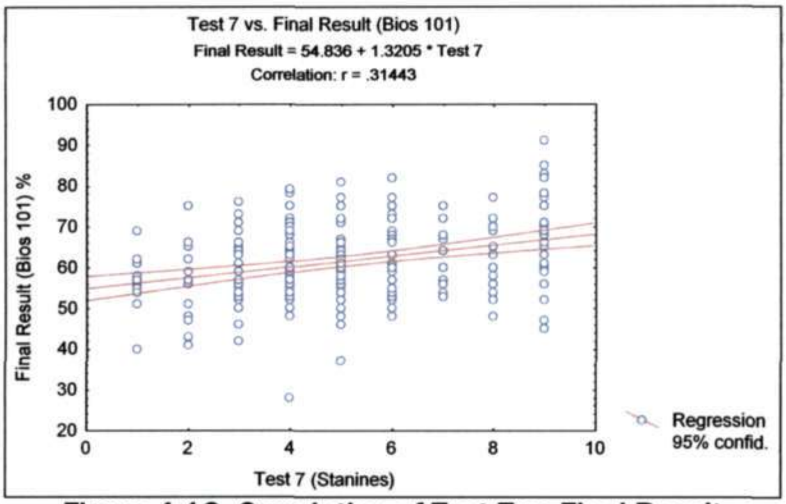


Figure 4.1G. Correlation of Test 7 vs Final Result

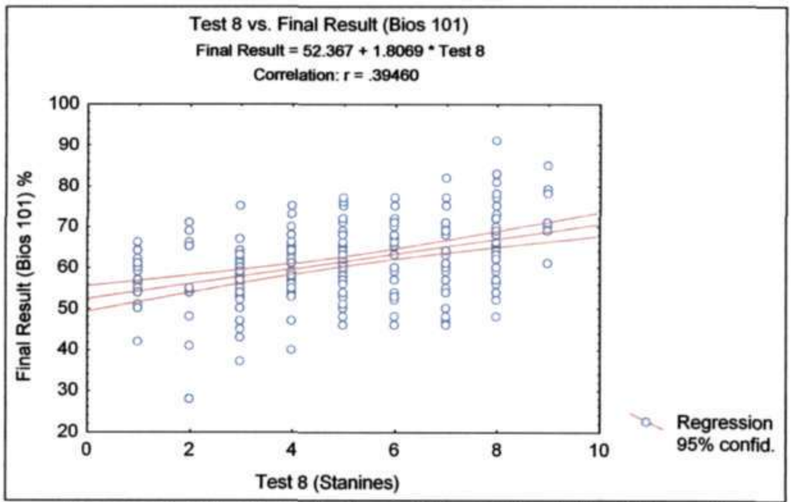


Figure 4.1H. Correlation of Test 8 vs Final Result

The correlation of the average of all tests and the final result is a reasonable r = .50396 (Figure 4.2A). The distribution of the average of all tests is also reasonable with the average slightly above the expected norm. The reverse is true for the final results in Biosciences for the 2004 cohort with the average slightly below the expected norm (Figure 4.2 B).

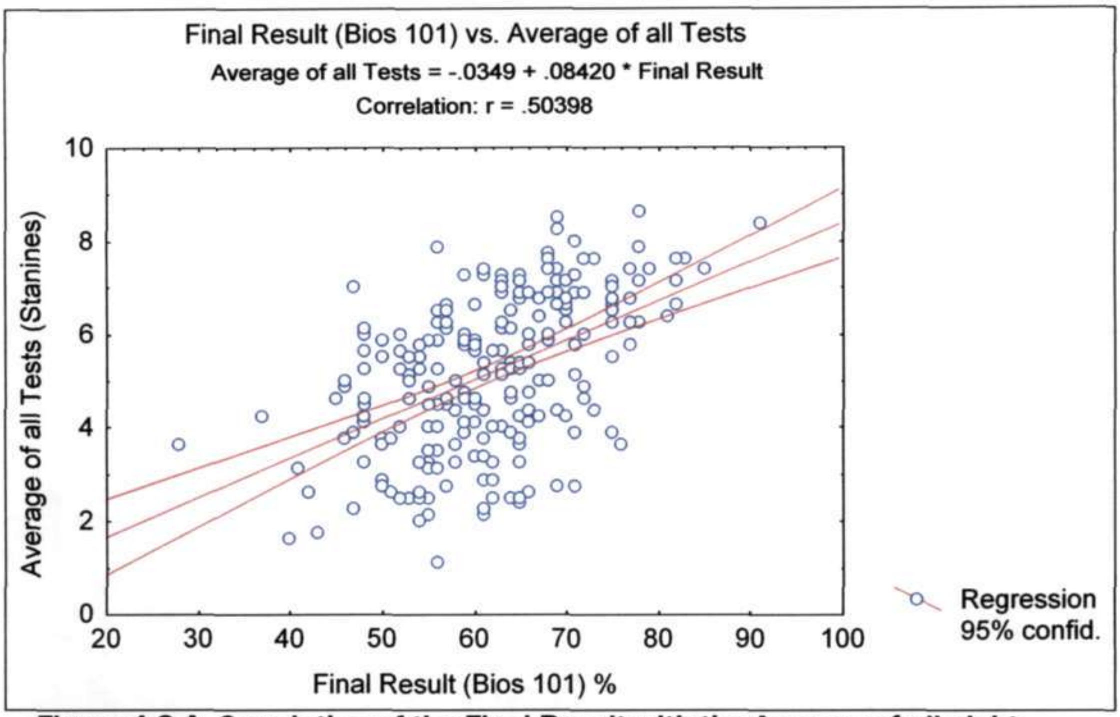


Figure 4.2 A. Correlation of the Final Result with the Average of all eight tests

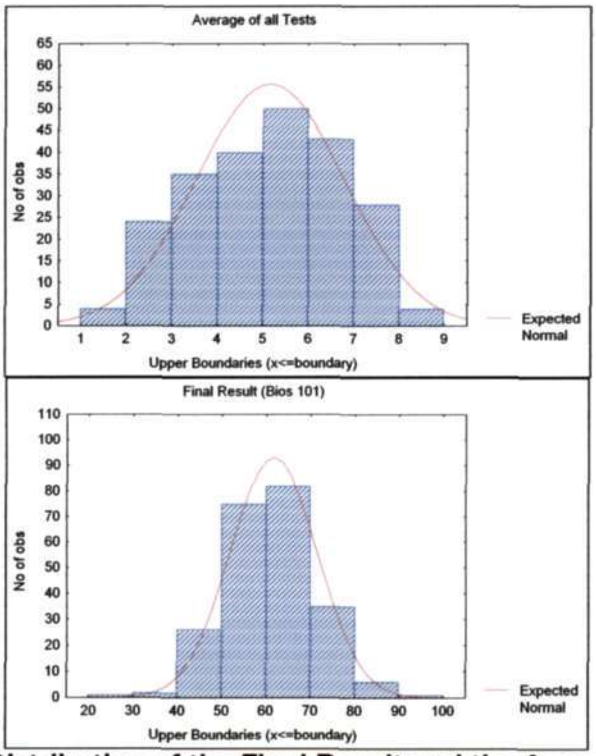


Figure 4.2 B. Distribution of the Final Result and the Average of all eight tests

The correlation of the final result with a combination of tests 2,3,4,5 is a reasonable r = .49462 (Figure 4.3 A) with the correlation of the final result with a combination of tests 4,5,6,7 is a weaker r = .45208 (Figure 4.3 B) and the correlation of the final result with a combination of tests 1, 2, 3,4,5,8 ("Academic" group of tests) is a good, r = .52224 (Figure 4.3 C) the best correlation of all of the DAT tests or combinations thereof. The distribution of the different combination of tests is reasonable although not completely comparable with the expected norm (Figure 4.3 D).

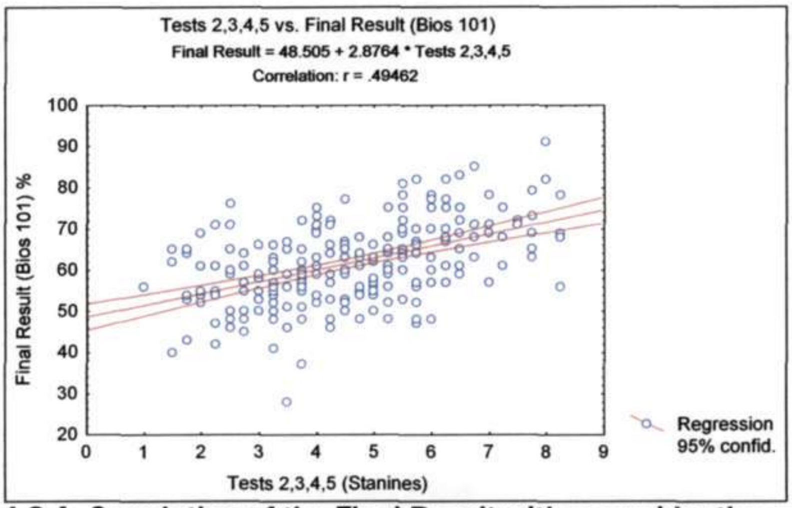


Figure 4.3 A. Correlation of the Final Result with a combination of Tests 2,3,4,5.

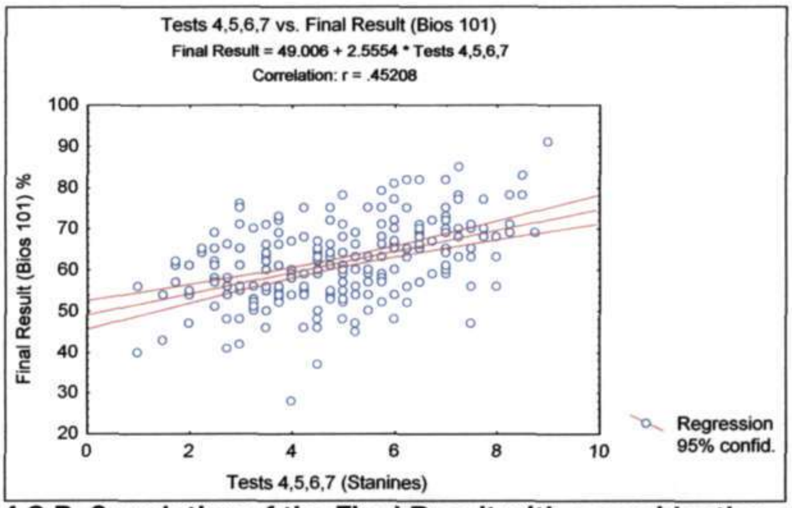


Figure 4.3 B. Correlation of the Final Result with a combination of Tests 4,5,6,7.

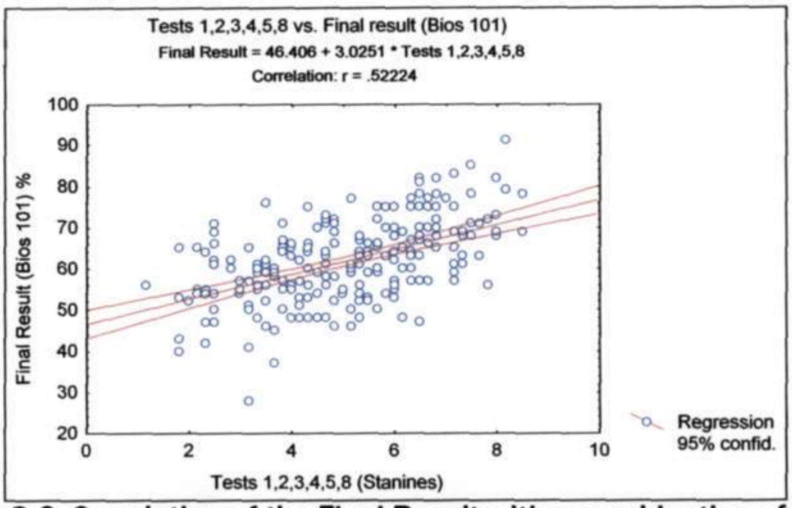


Figure 4.3 C. Correlation of the Final Result with a combination of Tests 1, 2,3,4,5,8.

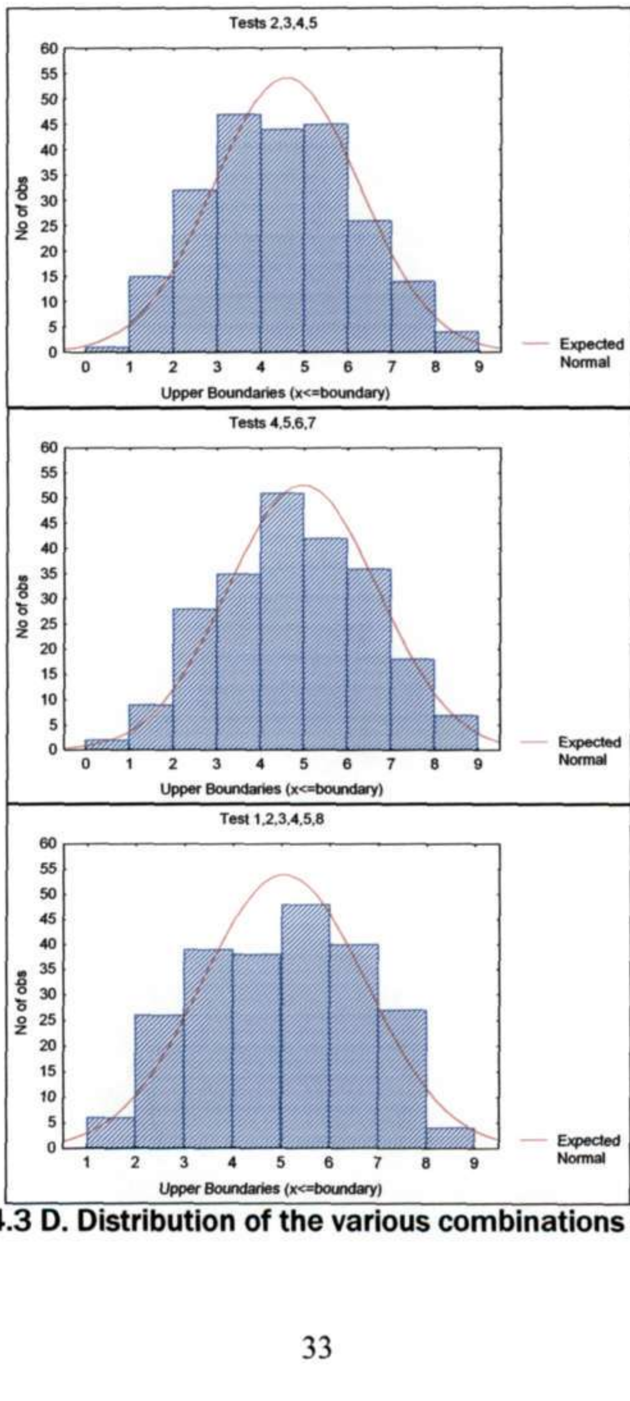


Figure 4.3 D. Distribution of the various combinations of tests.

4.3 Correlation of the DAT to other subjects.

The correlation with DAT and Biosciences is r = .50396 while the best correlation for the other subjects is Computer Science with r = .41165, with a declining correlation between Physics with r = .34085 followed by extremely poor correlations for Chemistry and Mathematics of r = .20313 and r = .08700 respectively (Figure 4.4 A-D). This indicates that when comparisons with the DAT and the different subjects, generally taken by first year non computational students, are undertaken, Biosciences shows a reasonable correlation while the other subjects show a declining, weaker correlation.

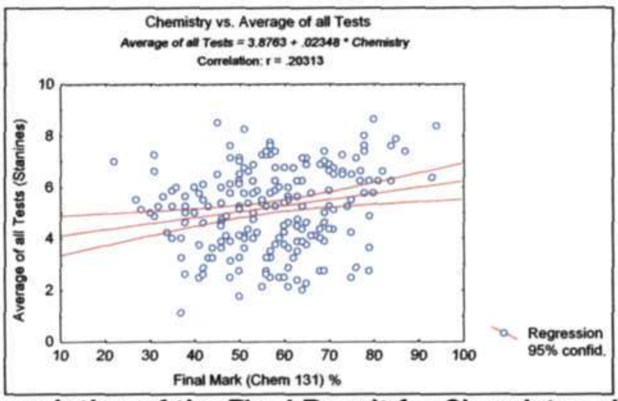


Figure 4.4 A. Correlation of the Final Result for Chemistry with the Average of all eight tests

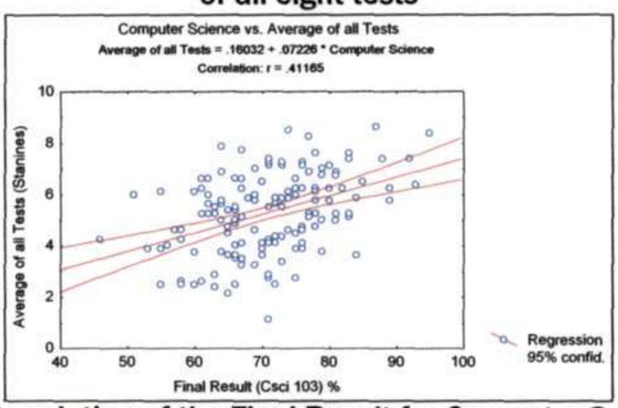


Figure 4.4 B. Correlation of the Final Result for Computer Science with the Average of all eight tests

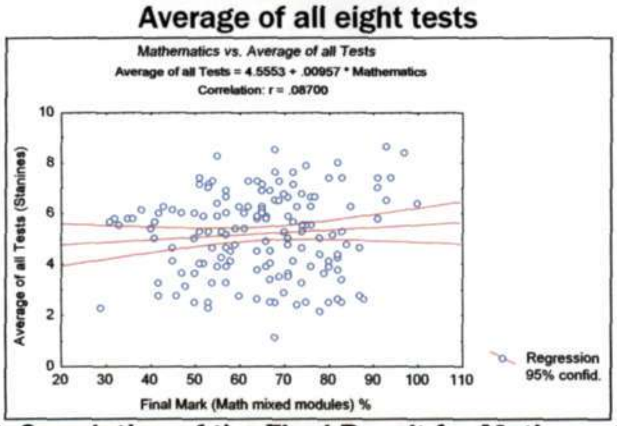


Figure 4.4 C. Correlation of the Final Result for Mathematics with the Average of all eight tests

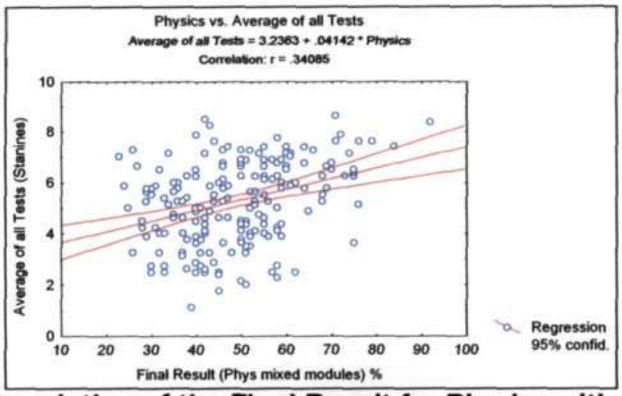


Figure 4.4 D. Correlation of the Final Result for Physics with the Average of all eight tests

4.4 Correlation of the DAT with the components that make up the Biosciences final mark.

The correlation with DAT and the Final mark for Biosciences is an acceptable r = .50396 but the correlations for the components that make up the final mark do not show as good a correlation. The correlation for the class mark (that is a form of formative assessment, made up of essays, tests and practical write ups) is r = .45844 while the correlation for the summative practical and theory exam is, r = .48146 and r = .42226 respectively. (Figures 4.5 A-C).

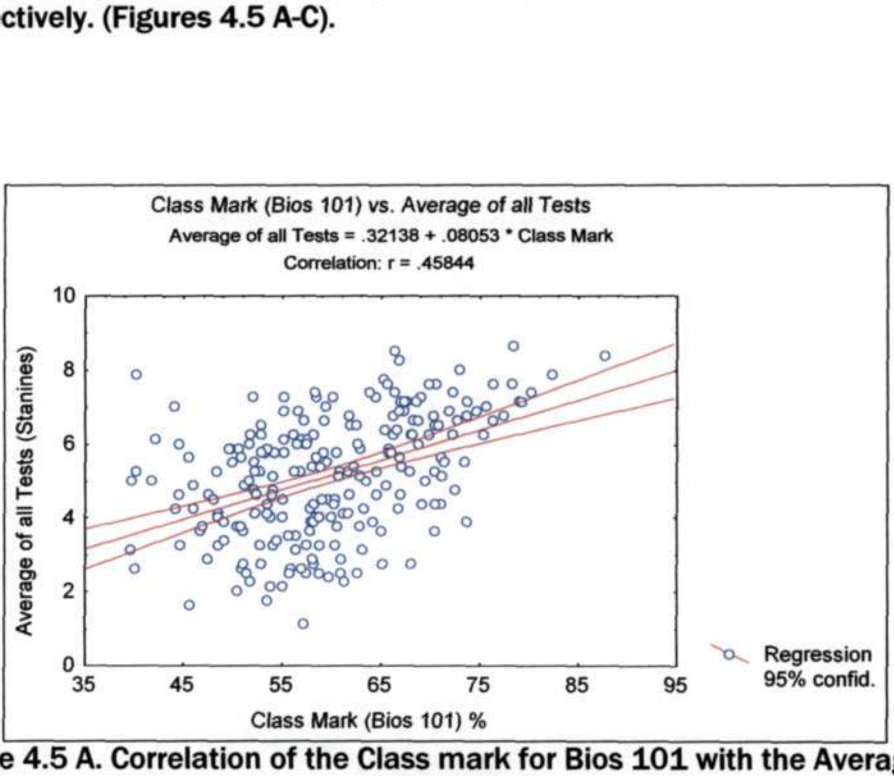


Figure 4.5 A. Correlation of the Class mark for Bios 101 with the Average of all eight tests

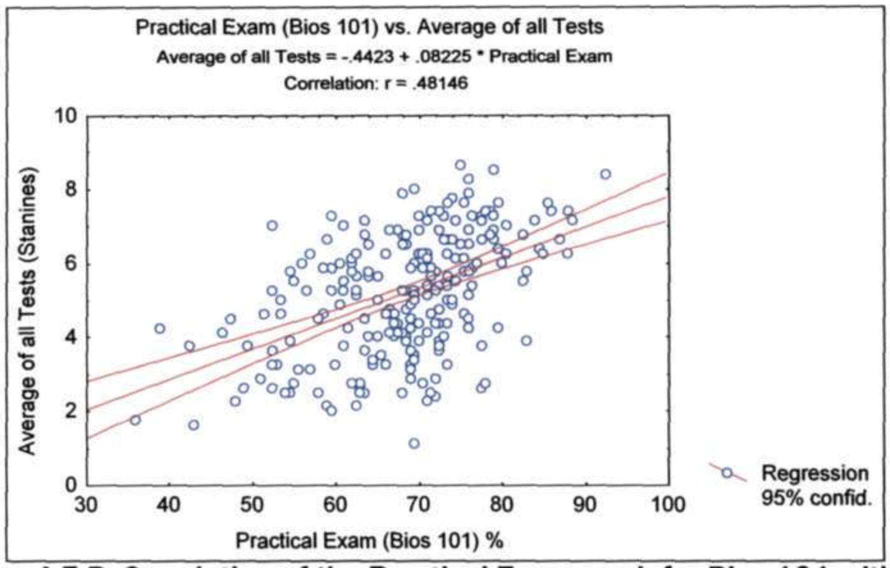


Figure 4.5 B. Correlation of the Practical Exam mark for Bios 101 with the Average of all eight tests

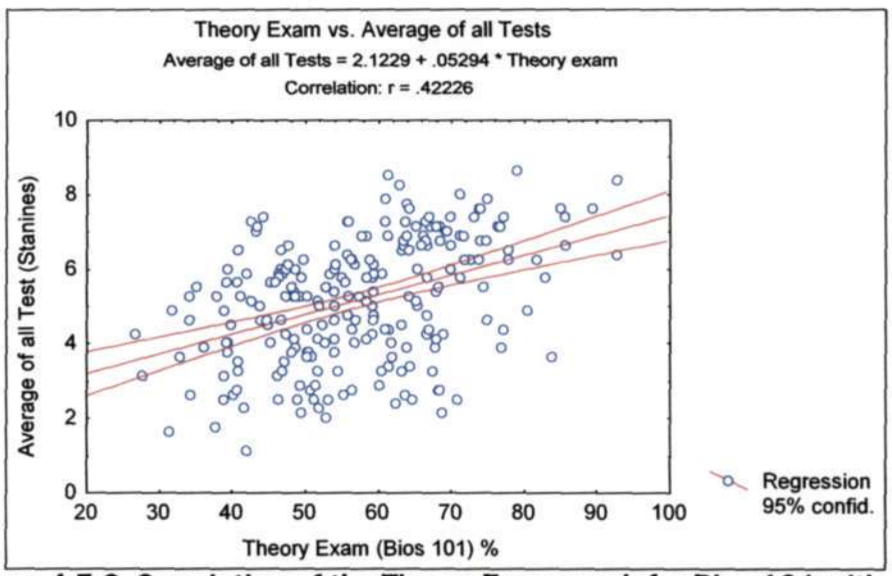


Figure 4.5 C. Correlation of the Theory Exam mark for Bios 101 with the Average of all eight tests

4.5 Correlation of the DAT and Final result (Bios 101) with the Matric points that the students obtained.

The best correlation that was obtained during this research was with the correlation between matric points and the average of all the DAT tests. The correlation with DAT and the Final mark for Biosciences is r = .50396 while that for DAT and the matric points is r = .57150. Interestingly, there is a poor correlation of r = .46732 if one compares the matric points with the Final mark obtained for Biosciences. Thus making the DAT a better predictor of student achievement than the matric points. The distribution of the matric points seem to be fairly similar to that of the expected norm (Figures 4.6 A-C).

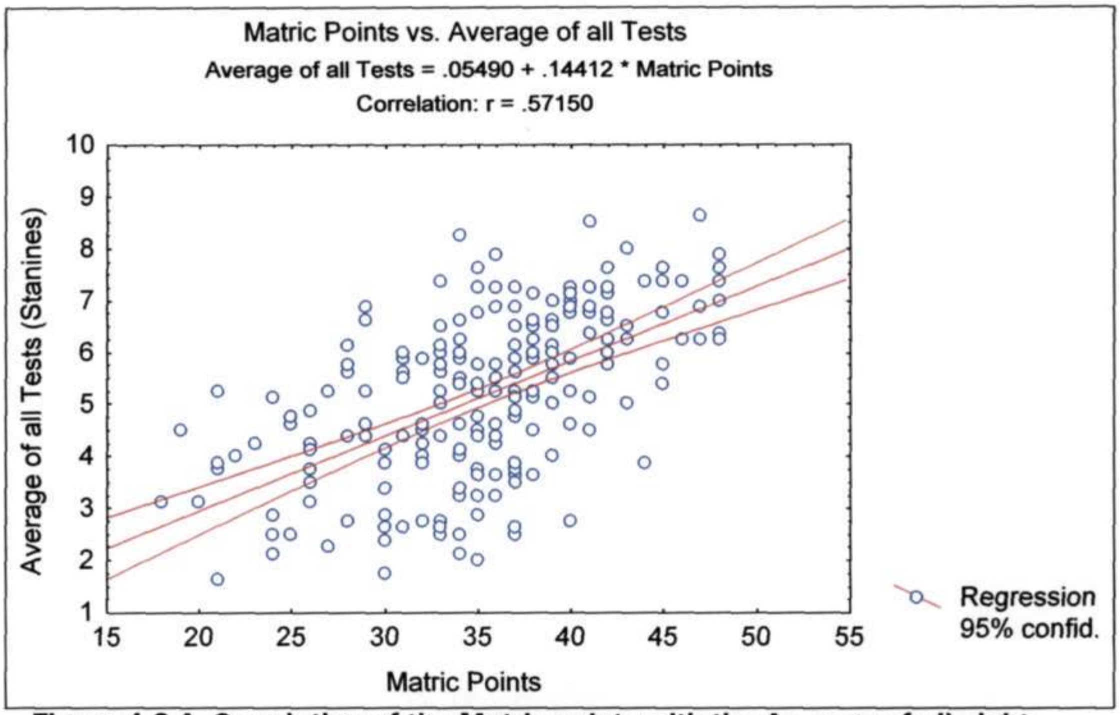


Figure 4.6 A. Correlation of the Matric points with the Average of all eight DAT tests

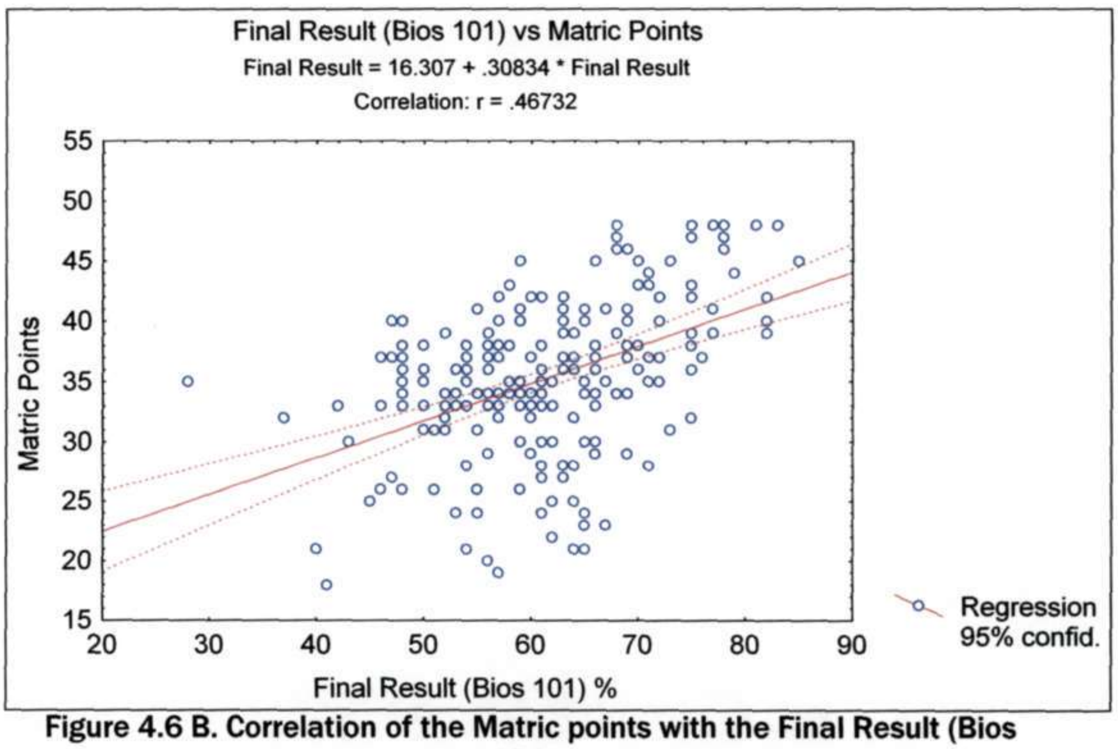


Figure 4.6 B. Correlation of the Matric points with the Final Result (Bios 101).

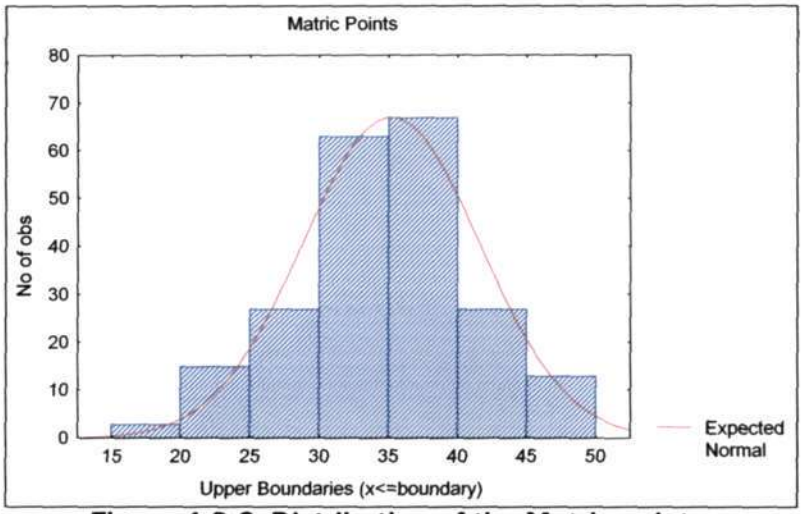


Figure 4.6 C. Distribution of the Matric points.

4.6 The effect of gender on DAT tests.

Table 4.2 and Table 4.3 outline the means of the various tests and variables that were recorded, using gender as the variable. Apart from test 3 and 4 (Non-Verbal Reasoning: Figures and Calculations) and the final results for Chemistry and Physics, females outperformed the males. Females attained a significantly higher average on the DAT tests but the difference between female and male on their final results in Biosciences is not significant (Table 4.4).

As evidence in Table 4.2 and Figure 4.7 females performed better in the three different combinations of the tests and did particularly well with the combination of test 4,5,6,7. The first series of tests indicates potential for further study while the middle series of test show clerical aptitude while the last series shows potential for higher education. Thus females are better suited in all these categories. Table 4.5 compares the components of the final result (class mark, exam mark and practical exam mark); the analysis shows no difference based on gender but rather differences based on test.

Table 4.2. The Means, Maxima, Minima and Standard Error showing the difference between the two genders.

	Male					Female				
	N	Mean (%)	Minimum (%)	Maximum (%)	Standard Error	N	Mean (%)	Minimum (%)	Maximum (%)	Standard Error
FINAL RESULT (BIOS 101)	102	60.81	40.00	91.00	0.90	126	62.44	28.00	83.00	0.91
CLASS MARK	102	58.88	40.20	87.80	0.89	126	61.09	39.70	82.50	0.84
THEORY EXAM MARK	102	56.04	31.50	92.80	1.23	126	58.49	26.80	92.80	1.20
PRAC EXAM MARK	102	67.93	36.00	92.50	0.93	125	68.44	39.00	88.00	0.87
TEST 1	102	6.45	1.00	9.00	0.26	126	7.12	1.00	9.00	0.19
TEST 2	102	5.72	1.00	9.00	0.21	126	6.11	2.00	9.00	0.18
TEST 3	102	3.58	1.00	8.00	0.19	126	3.40	1.00	9.00	0.18
TEST 4	102	5.11	1.00	9.00	0.21	126	4.96	1.00	9.00	0.17
TEST 5	102	3.56	1.00	9.00	0.22	126	4.22	1.00	9.00	0.19
TEST 6	102	5.06	1.00	9.00	0.23	126	6.27	1.00	9.00	0.18
TEST 7	102	4.84	1.00	9.00	0.24	126	5.50	1.00	9.00	0.20
TEST 8	102	5.04	1.00	9.00	0.21	126	5.28	1.00	9.00	0.19
AVE OF ALL TESTS	102	4.92	1.13	8.38	0.17	126	5.36	2.38	8.63	0.14
TESTS 2,3,4,5	102	4.49	1.00	8.25	0.17	126	4.67	1.50	8.25	0.15
TESTS 4,5,6,7	102	4.64	1.00	9.00	0.18	126	5.24	2.00	8.75	0.14
TESTS 1,2,3,4,5,8	102	4.91	1.17	8.17	0.18	126	5.18	1.83	8.50	0.14
CHEMISTRY FINAL MARK (CHEM131)	92	58.71	28.00	94.00	1.35	114	55.85	22.00	93.00	1.39
COMPUTER SCIENCES FINAL MARK (CSCI 103)	65	69.98	46.00	95.00	1.05	76	71.82	51.00	93.00	1.05
MATHS (MIXED MODULES)	74	64.20	29.00	97.00	1.67	81	65.22	31.00	100.00	1.70
PHYSICS (MIXED MODULES)	80	49.58	29.00	92.00	1.39	110	48.23	23.00	79.00	1.33

Table 4.3. Effect of Gender on Average of all Tests

Effect of Gender on Average of all Tests									
Analysis of Variance									
Marked effects are significant at p < .05000									
	SS Effect	df Effect	MS Effect	SS Error	df Error	MS Error	F	p	
AVE	10.815	1	10.815	593.355	226	2.625	4.119	0.043	<u>***SD***</u>

Scheffe Test; Variable:			
Marked differences are significant at p < .05000			
	{1}	{2}	
	M=5.357	M=4.919	
F {1}		<u>*0.0436*</u>	
M {2}	<u>*0.0436*</u>		<u>***Females attained a higher average on the tests***</u>

Table 4.4. Effect of Gender on Final Result (Bios 101)

Effect of Gender on Final Result (Bios 101)								
Analysis of Variance								
Marked effects are significant at p < .05000				***No difference in Gender***				
	SS	df	MS	SS	df	MS		
	Effect	Effect	Effect	Error	Error	Error	F	p
FINALRES	148.442	1	148.442	21496.453	226	95.117	1.561	0.213

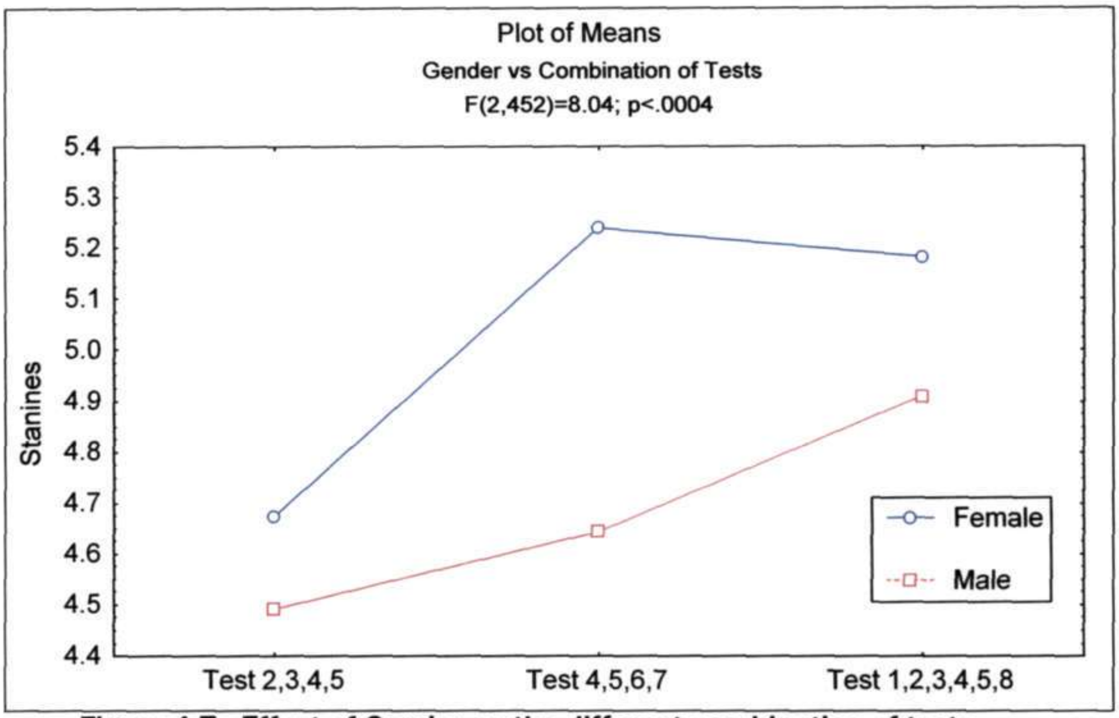


Figure 4.7. Effect of Gender on the different combination of tests.

Table 4.5. Effect of Gender on Class Mark & Practical & Theory Exam

Effect of Gender on Class Mark & Practical & Theory Exam						
Summary of all Effects				***No difference because of gender but because of component***		
1-GENDER, 2-RFACTOR1						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	556.031	225	258.182	2.154	0.144
2	2	7145.836	450	42.597	167.755	<u>0</u>
12	2	73.530	450	42.597	1.726	0.179

4.7 The effect of language on DAT tests.

When home language was compared to the average of all the DAT tests and the final result for Biosciences (Figure 4.8 and 4.9 and Table 4.6 and 4.7) it was found that language has a profound effect on these results.

Interestingly,[5] students who have English as a home language did not perform the best in the DAT tests nor the Bioscience course. The language designated "?" is a combination of home languages from non South African students. These students perform particularly badly. Apart from the students that are designated "Swati" the trend in the DAT tests generally mirrors their performance in the course as a whole. Because of the small number of samples for some of the languages, the results were not statistically calculated for each language. The graphs primarily show the trends that are observed for the different languages; although the box plot (Figure 4.9 does attempt to put some form of stats on these trends). Rather the calculations were done to see if language in general was significant in the attainment of the results. The designation "Graduate" reflects the results from graduate student demonstrators who were assisting at the time the tests were given and completed the test out of interest.

Home Languages recorded for the tests.

E	English
?	Unknown
Z	Zulu
S	Setswana
T	Tsonga
A	Afrikaans
SW	Swati
X	Xhosa
SS	South Sotho
G	Graduate
NS	North Sotho
V	Tshvenda
N	Ndebele

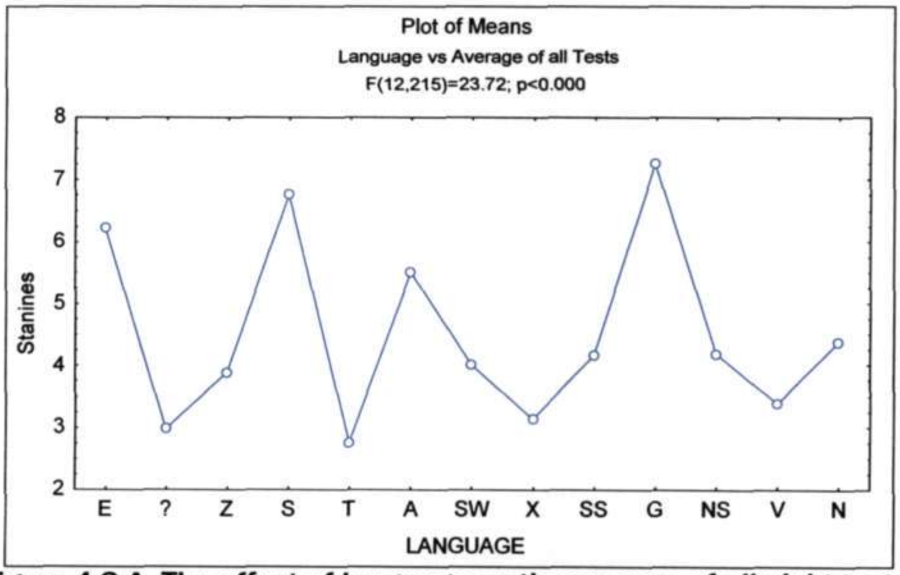


Figure 4.8.A. The effect of language on the average of all eight tests.

[5] English is the language of instruction for the Bioscience course and the language that the DAT tests were conducted in.

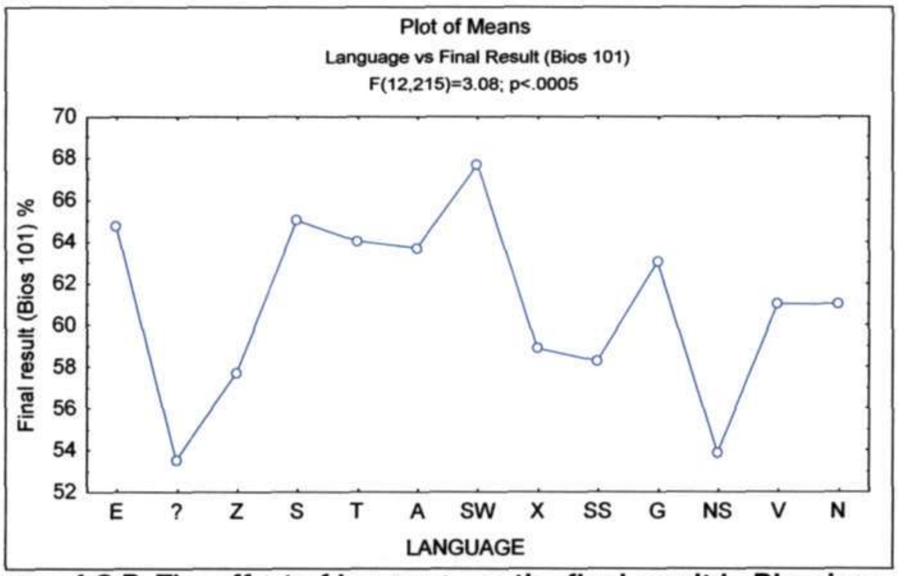


Figure 4.8.B. The effect of language on the final result in Biosciences.

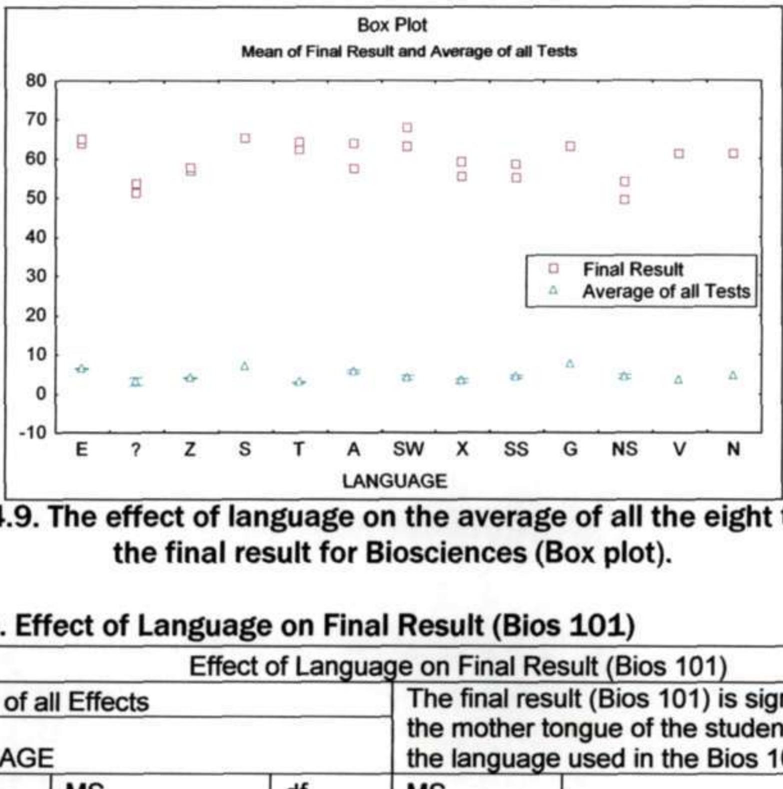


Figure 4.9. The effect of language on the average of all the eight test and the final result for Biosciences (Box plot).

Table 4.6. Effect of Language on Final Result (Bios 101)

Effect of Language on Final Result (Bios 101)						
Summary of all Effects 1-LANGUAGE				The final result (Bios 101) is significantly affected by the mother tongue of the student when it differs from the language used in the Bios 101 course.		
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	12	264.391	215	85.917	3.077	***0.0005***

Table 4.7. Effect of Language on the Average of all Tests

Effect of Language on the Average of all Tests						
Summary of all Effects 1-LANGUAGE				The average of all the tests is significantly affected by the mother tongue of the student when their language differs from that used in the DAT-L test		
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	12	28.68079	215	1.209306	23.71673	***0***

4.8 The effect of race groups on DAT tests.

Perhaps most contentious of all, is the question of whether race plays a role in performance in the DAT's and in the exams. Table 4.8 outlines the means of the various tests and variables that were recorded and compares these against race groups. The students designated coloured by the SMS have not been included in the table due to their small number. Tables 4.9- 4.11 and Figure 4.10 summarize the performances of the different race groups against the different variables. White students perform significantly better than the other race designations when it comes to the final result. Within the categories African, Indian and Coloured[6] (there is no significant difference (Table 4.9). When it comes to the different DAT tests, Africans perform significantly worse than the other three groupings and the Indian grouping did significantly worse when compared to Whites. The rest of the comparisons were not significant (Figure 4.10 and Table 4.10). When performing a repeated measure RMANOVA; race group, the different tests and the combination of both are all significantly different (Table 4.11).

Figure 4.11 and Table 4.12 both show that race impacts on marks obtained in the class marks, exam marks and the theory exam mark. Race and the different components of the final exam, when taken as single entities are definitely significantly different, while the combination of the two together is not significant.

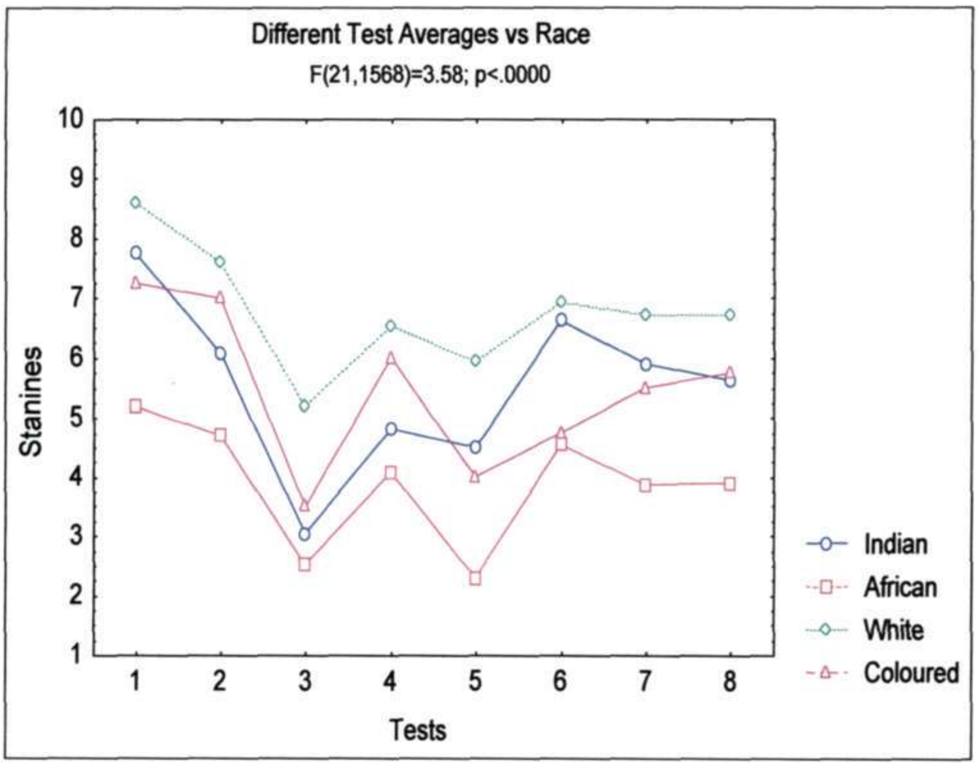


Figure 4.10. The effect of race on the score attained for the different tests.

[6] The Coloured group has been included for comparison purposes only.

Table 4.8 **The Means, Maxima, Minima and Standard Error showing the difference between races.**

	White					African					Indian				
	N	Mean (%)	Minimum (%)	Maximum (%)	Standard Error	N	Mean (%)	Minimum (%)	Maximum (%)	Standard Error	N	Mean (%)	Minimum (%)	Maximum (%)	Standard Error
FINAL RESULT (BIOS 101)	71	69.66	48.00	91.00	0.91	105	58.38	28.00	77.00	0.84	48	57.71	45.00	81.00	1.20
CLASS MARK	71	67.56	40.30	87.80	1.01	105	57.27	39.70	73.80	0.72	48	55.78	39.80	80.30	1.14
THEORY EXAM MARK	71	66.37	42.80	92.80	1.29	105	53.66	26.80	84.00	1.18	48	52.95	34.30	92.80	1.71
PRAC EXAM MARK	71	74.99	58.50	92.50	0.81	104	65.33	36.00	83.00	0.90	48	64.71	46.50	85.00	1.32
TEST 1	71	8.58	6.00	9.00	0.08	105	5.19	1.00	9.00	0.23	48	7.75	1.00	9.00	0.23
TEST 2	71	7.59	3.00	9.00	0.18	105	4.70	1.00	9.00	0.16	48	6.08	2.00	9.00	0.27
TEST 3	71	5.20	1.00	9.00	0.24	105	2.51	1.00	6.00	0.13	48	3.04	1.00	8.00	0.20
TEST 4	71	6.52	1.00	9.00	0.20	105	4.08	1.00	8.00	0.17	48	4.81	1.00	8.00	0.25
TEST 5	71	5.94	1.00	9.00	0.20	105	2.30	1.00	6.00	0.12	48	4.50	1.00	9.00	0.28
TEST 6	71	6.93	2.00	9.00	0.19	105	4.54	1.00	9.00	0.21	48	6.63	2.00	9.00	0.27
TEST 7	71	6.70	3.00	9.00	0.23	105	3.87	1.00	9.00	0.19	48	5.90	2.00	9.00	0.28
TEST 8	71	6.72	3.00	9.00	0.19	105	3.90	1.00	8.00	0.17	48	5.63	1.00	9.00	0.26
AVE OF ALL TESTS	71	6.77	4.75	8.63	0.10	105	3.89	1.13	6.75	0.12	48	5.54	2.75	7.38	0.14
TESTS 2,3,4,5	71	6.31	3.25	8.25	0.13	105	3.40	1.00	6.25	0.11	48	4.61	2.50	7.25	0.16
TESTS 4,5,6,7	71	6.52	4.25	9.00	0.13	105	3.70	1.00	7.00	0.12	48	5.46	3.00	8.00	0.17
TESTS 1,2,3,4,5,8	71	6.76	4.33	8.50	0.11	105	3.78	1.17	6.33	0.12	48	5.30	2.50	7.33	0.15
CHEMISTRY FINAL MARK (CHEM131)	65	62.98	31.00	94.00	1.73	93	56.58	27.00	80.00	1.24	44	50.55	22.00	93.00	2.24
COMPUTER SCIENCES FINAL MARK (CSCI 103)	47	74.66	60.00	95.00	1.10	60	67.58	46.00	89.00	1.08	30	72.73	57.00	93.00	1.73
MATHS (MIXED MODULES)	50	66.84	31.00	97.00	2.05	70	66.93	29.00	91.00	1.63	31	57.16	33.00	100.00	2.83
PHYSICS (MIXED MODULES)	62	56.82	26.00	92.00	1.76	85	46.41	26.00	76.00	1.16	40	41.98	23.00	75.00	1.91

Table 4.9. The effect of race on the final result for Biosciences 101

ANOVA	Final Result (Bios 101) vs Race						
Summary of all Effects							
1-RACE							
	df	MS	df	MS			
	Effect	Effect	Error	Error	F	p-level	
1	3	2184.11	224	67.378	32.416	***1.93E-17***	**SD**

Probabilities for Post Hoc Tests Scheffe test				
MAIN EFFECT: RACE				
	I	A	W	C
	57.708	58.381	69.662	56.000
I		0.974	***1.2E-11***	0.983
A	0.974		***8.69E-15***	0.955
W	***1.2E-11***	***8.69E-15***		***0.016***
C	0.983	0.955	***0.016***	

Table 4.10. The effect of race on the different DAT tests

ANOVA	Different Tests vs Race						
Summary of all Effects							
1-RACE							
	df	MS	df	MS			
	Effect	Effect	Error	Error	F	p-level	
1	3	120.857	224	1.079	112.05	***0***	**SD**

Probabilities for Post Hoc Tests Scheffe test				
MAIN EFFECT: RACE				
	I	A	W	C
	5.542	3.886	6.773	5.469
I		***2.24E-15***	***4.39E-08***	0.999
A	***2.24E-15***		***0***	***0.032***
W	***4.39E-08***	***0***		0.116
C	0.999	***0.032***	0.116	

Table 4.11. The effect of race on the different DAT tests.

RMANOVA	Different Tests vs Race						
Summary of all Effects							
1-RACE, 2-TESTS							
	df	MS	df	MS			
	Effect	Effect	Error	Error	F	p-level	
race	3	966.859	224	8.628	112.054	***0***	SD
tests	7	70.405	1568	2.321	30.338	***0***	SD
race + test	21	8.315	1568	2.321	3.583	***8.1E-08***	SD

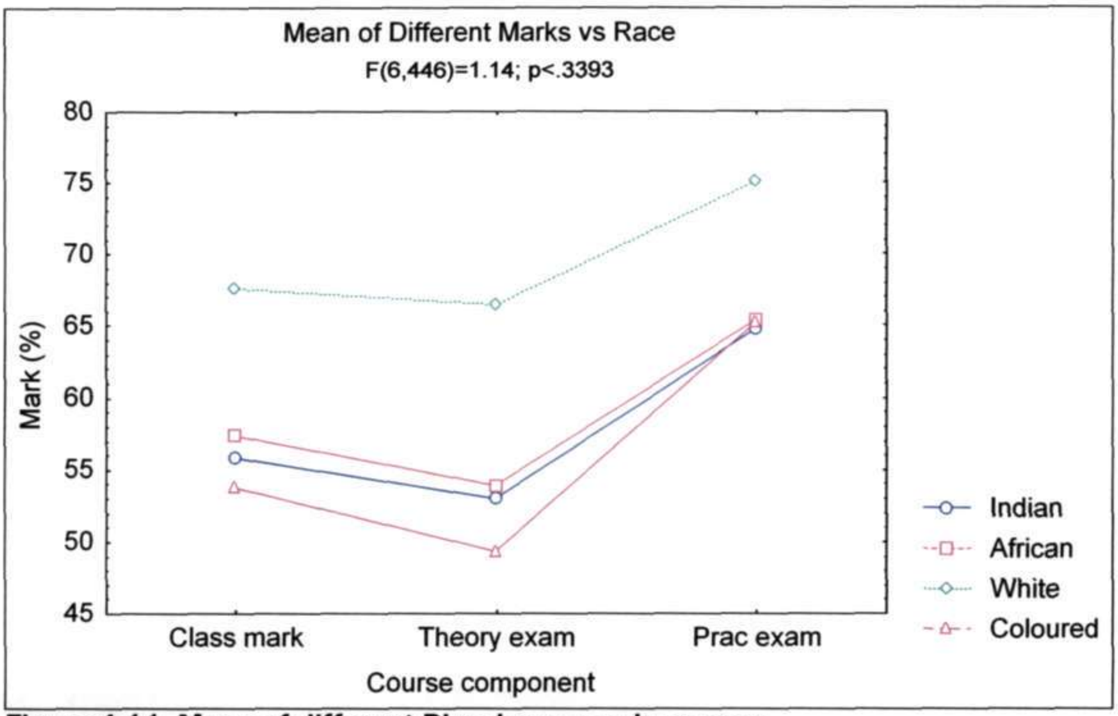


Figure 4.11. Mean of different Bioscience marks vs race.

Table 4.12. The effect of race on the different components of the final mark for Biosciences 101.

RMANOVA				Prac & Theory Exam & Class Mark vs Race		
Summary of all Effects						
1-RACE, 2-RFACTOR1						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	3	6167.43	223	180.021	34.260	***2.99E-18***
2	2	2089.037	446	42.656	48.975	***5.92E-20***
12	6	48.531	446	42.656	1.138	0.339

4.9The effect of SFP on DAT tests.

Tables 4.13 and 4.14, highlight the fact that although students who have completed SFP did not show a significant difference (against the rest of the class) when it came to the final mark for Bioscience, there was however a significant difference on their result s for the different DAT tests.

Table 4.13. The effect of SFP on the average of all the tests.

Effect of SFP on the Average of all Tests						
Summary of all Effects				***SFP did significantly worse than rest of the class with reference to the average of all the tests***		
1-SFP						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	1	77.34607	225	2.322933	33.29673	***2.6E-08***

Table 4.14. The effect of SFP on the final result for Biosciences 101.

Effect of SFP on the Final Result (Bios 101)						
Summary of all Effects				***SFP were not significantly different from the rest of the class with reference to the final result***		
1-SFP						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	1	267.725	225	93.095	2.876	***0.091***

5 Discussion

The discussion will attempt to tease out to what extent aptitude tests are fair predictors of university performance and the effects aptitude testing may have on the education system. Because controversy has surrounded group differences in test scores, a discussion will determine whether these differences are an accurate reflection of students' academic potential or result from limitations of the tests.

Aptitude tests provide information on a learner's relatively strong and weak aptitudes. Based on this information deductions and predictions can be made about possible success in specific fields of study. There has been some discussion about the accuracy of the norms provided for the DAT-L, but the data produced by this study is a starting point.

With respect to the DAT-L it is important to note that the norms furnished for the DAT-L and used in this research are subject to serious shortcomings. The sample used to construct the norms tables is not an accurately reflection of the distribution of race in South Africa, and Owen *et al.* (2000) admit that there was not sufficient representation among black students across the educational spectrum. Because of the problems experienced during the standardization of the DAT-L, the norms should be regarded only as a broad guideline.

Owen *et al.* (2000) state that the single test that gives the best indication of the general intellectual level of the learner is Verbal Reasoning. However, a more reliable indication will be obtained by taking into consideration also the scores for Non-verbal Reasoning[7], Calculations and Reading Comprehension. If the learner obtains a mean stanine of 7 or higher in these four tests together, he/she will fall in the top 23 percent of the population and can be regarded as university material! The average for all these tests for the Bioscience 2004 cohort was 4.59.[8] This indicates that on average our first year Biosciences students are 'average'. The capacity to think logically and critically, to be able to order ideas and construct an argument both verbally and in writing is essential in biology and the lack of this capacity must be addressed in our 'average' students.

The findings of this study show that the average for the White group is 6.31, while the figures are 3.4 and 4.61 for the African and Indian groups respectively. Across the board the potential of the student in this cohort is average or below average, given that the norms are calculated for the general population. Yet the result the students attained as their final results mask this! This raises two questions, the answers to which are beyond the scope of this research. Firstly, are we accepting students that should not be at university? Secondly, have we dropped our standards to accommodate

[7] Especially in the case of learners with a language backlog.

[8] A mean stanine of 3 or lower means that the student falls in the bottom 23 percent of the population and will probably find it difficult to progress beyond the school grade in which they are currently in; with a mean stanine of 4, 5 or 6 the candidate falls between the 23rd and 78th percentile, and can be regarded as an average learner.

these average students?[9]

A similar trend is evident in all the combinations of aptitudes, whether these test for academic or clerical potential.

According to Owen *et al.* (2000) the tests for verbal reasoning, reading comprehension, memory etc concerning verbal material, and so on, cannot be measured if there is a language backlog[10]. The results from the vocabulary test are reasonable across the board (6.82) so these vocabulary results cannot entirely be used to justify the poor performance of the students in the other results. However, language is often regarded as an important moderator when analysing test performance. Poor performance on assessment measures could be more a reflection of language difficulties than of weak individual ability as it may take longer to process information if English is not the home language[11].

The Vocabulary test is regarded by Owen *et al.* (2000) as a valid indication of a students' knowledge of the meaning of words and as a valid criterion for the Verbal Comprehension factor. In the majority of the cases this test was relatively well done. The test for Verbal Reasoning, which assumes that the student has the ability to determine relationships and to complete word analogies, to solve general problems requiring logical thought was also reasonably completed. However, the test for Non-verbal Reasoning (which tests the ability to see the relationships between figures and, by analogy, to identify an appropriate missing figure, as well as following the changes that the figures of a figure series undergo, to deduce the work principle and to apply it again), was very poor for all groups across the board and offers the greatest challenge to teachers at the first year level in addressing this. Calculations were reasonable.

The students' capacity to complete the Reading Comprehension test was poor. Reading Comprehension is claimed to be an important indicator which should tie in with the results from the Vocabulary and Verbal Reasoning tests which it did not. The Reading Comprehension test aims to measure the learner's ability to comprehend what he /she is reading. The choice of passages is perhaps too advanced for some of the first year students and the contexts of the readings were very Eurocentric. The time available to the students may also have contributed to the poor performance.

The Comparison test (which assesses visual perceptual speed) was adequate. Price Controlling was not as badly answered, as was the memory test. This result for the memory test was surprising as there is the perception that a number of students in this group relies heavily on rote

[9] It is important to note that the assessment in the Biosciences course is norm referenced as opposed to criterion referenced.

[10] See later discussion on language.

[11] However, one cannot separate linguistic competence from cognitive competence. Research shows that the two variables are too integrated, we think in language. If a student is operating in a second language and is not fully bilingual, then it must affect cognitive performance negatively.

learning, which in turn relies heavily on memory without necessarily understanding the subject matter!

Based on the description of the stanine scale; the results for the White students place them in the range above average to good while the Indian students were slightly below average, but the African students were in the poor to below average range. This result has serious implications for our approach to teaching the first years[12]. Although these results based on the DAT tests are alarming, the result that students obtained in the final exam does not paint as bleak a picture.

We would like to think that the improved performance in first year is a result of they way we teach and assess skills in the first semester. The skills that we attempt to teach and develop in our students during the first semester are clearly outlined to the students before the start of the course and are reinforced throughout the course. Our assessment is primarily formal assessment; however, it is both of a formative and summative nature, allowing the students to help themselves to learn and to allow them to measure how much they have learnt. Thus we have built into the system a feedback device, and with a summative system placed at the end. There is a good deal of continuous assessment as we do not go onto new skills after a test / practical, but continue to integrate and develop the previous skills and give feedback and guidance on them. The way we construct our essays and questions is to try and encourage a 'deep approach' to learning. We try to maximise understanding, and encourage the student to read widely, to discuss issues and reflect on what has been heard and read. The way some of our questions are framed is to encourage integrating details into broad, over-arching ideas which students are hopefully constantly trying to develop and as a result understand learning to involve meaning, understanding, and ways of interpreting the world. The ways we have chosen to assess our students also influences their styles of learning. In the past the use of multiple choice questions and other forms of tests promoted reproductive styles of learning whereas more recently projects and open-ended assessments have promoted independence and deeper strategies of understanding.

The alternative scenario is that we have dropped our standards to accommodate the average student. Although, there has been a slight lowering of standards, we would like to think that this has been minimal as a result of the introduction of the developmental and scaffolding strategies that we have introduced into our first year Biosciences course.

The sequence of correlation for the different tests against the final Bioscience result is from Vocabulary through Reading Comprehension, Verbal Reasoning [Mellanby *et al.* (2000) show that verbal reasoning test scores was a moderate predictor of degree class], Calculation, Memory, Price Controlling, Non-Verbal Reasoning: Figures, and finally Comparisons. The sequence mostly makes sense, except the poor correlation with Non-Verbal Reasoning: Figures. The poor performance of the students in the

[12] To be discussed later.

Non-Verbal Reasoning test needs further investigation and this includes investigation into the way this is taught at the school level and in our first year.

Some vital skills that are essential for a biologist are the ability to: distinguish inferences from observations, make appropriate approximations; estimate; and have a sense of orders of magnitude, interpret abstract representations e.g. graphs, and translate between biological phenomena and representations e.g. diagrams. The poor Non-Verbal Reasoning skills of the student impacts directly on these skills and thus it is essential that developmental strategies are taken to ensure that the students can improve their capabilities in these skills.

The poor results from the Reading Comprehension tests indicate the need to focus on scaffolding strategies in this area. If students do not have the capacity to comprehend what they are reading then their capacity to perform well at University is significantly undermined. The reading and writing course that was offered to first years in the past should be reinstituted. Any student with below average reading and comprehension skills should be obliged to take this course and learn or refine these skills in addition to developmental changes in the curriculum. However, the development of these skills must be within the context of the discipline i.e. using readings from the course itself.

In order for DAT test scores to be used in guidance for selection, it is necessary that they have some predictive validity. McDonald *et al.* (2000) report that, the association between the SAT and GPA in the USA varied between colleges from a high of 0.72 to a low of 0.37.[13] Results from the current research shows that the correlation of the average of all tests and the final Bioscience result is a reasonable r = .50396 (Figure 4.2A). The correlation of the final result with a combination of tests 2,3,4,5[14] is also a reasonable r = .49462 (Figure 4.3 A) whilst the correlation of the final result with a combination of tests 4,5,6,7[15] is a weaker r = .45208 (Figure 4.3 B) and the correlation of the final result with a combination of tests 1, 2,3,4,5,8[16] is a good, r = .52224 (Figure 4.3 C), the best correlation of all of the DAT tests or combinations thereof.

What are the implications of these findings for the South African higher education context? I suggest that these correlations and the fact that weaker correlations are used in the USA as a means of predicting potential mean that an immediate response chould be that the DAT-L be implemented as a mechanism for the selection of students at the University

---

[13] In Israel Beller (1994) found the PET to be a good predictor of FGPA, with the average correlations being 0.53 for liberal arts, 0.5 for sciences, 0.45 for social sciences, and 0.43 for engineering. In the United Kingdom, the mean A-level grade was the best predictor of degree performance followed by school assessment with average correlations of 0.36 and 0.26 respectively.

[14] Tests 2,3,4,5: Verbal reasoning, Non-Verbal reasoning :Figures, Calculations, Reading Comprehension

[15] Tests 4,5,6,7:Calculations, Reading Comprehension, Comparisons, Price Controlling

[16] Tests 1,2,3,4,5,8 Vocabulary, Verbal reasoning, Non-Verbal reasoning :Figures, Calculations, Reading Comprehension, Memory

of KwaZulu-Natal. This is reinforced by the fact that tests 1,2,3,4,5,8 (which gives a good correlation) has a reduced number of tests and is relatively easy to administer.

However, these correlations are for the Biosciences and the sample size is fairly small. Despite this, the test bank shows promise for further research. One must however note the following precaution by McDonald *et al.* (2000) that, in the USA, both aptitude tests and high school records are able to predict performance to a modest degree. High school record has generally been found to be a better predictor than the SAT, although some studies suggest each predicts college grades to approximately the same degree.

Chapman (1996) found high positive correlations for a number of subjects; but the strongest links were seen for biology (0.47) while other commonly studied subjects showed poor correlation. The results from the current research show that correlation with DAT and Biosciences is r = .50396 while the best correlation for the other subjects is Computer Science with r = .41165, with a declining correlation for Physics r = .34085 followed by extremely poor correlations for Chemistry and Mathematics of r = .20313 and r = .08700 respectively (Figure 4.4 A-D).

The trend where biology shows a good correlation with high school marks or aptitude tests appears to be world wide. A major problem however, is that if we were to use the DAT as a means of selecting students it would be very effective for biology but not necessarily as effective for the other basic science subjects. The fact that the tests would probably have to be administered to all applicants who apply for admission to first year in the Science and Agriculture Faculty, makes the choice of this set of tests questionable. However, should the Faculty require a means of determining entrance into biologically orientated subjects, then the tests show great promise.

The correlation of the DAT with the individual components that make up the Biosciences final mark showed that a combination of all three components (from both summative and formative types of assessment[17]) gave the best value for comparison purposes.

The debate about the relevance of matric results is ongoing. Behr (1985), Bokhorst *et al.* (1992), and Van Wyk and Crawford (1984) all found that that successful first-year students obtain significantly better (higher) matric symbols than those of unsuccessful students. This was also found in this research but to a lesser degree. However, Botha and Cilliers (1999) have shown that matric symbols are a poor predictor of academic success. When comparing the matric points with the DAT and the final mark, there was a better correlation with the DAT tests than with the final mark. What this highlights is the fact that in this research, which was primarily focussed at biology students, the use of the results from the aptitude tests gives a better indication (predictor) of the performance of the student in Biosciences than if the matric points were used to predict performance.

---

[17] See previous comments.

This research suggests there is a good correlation between high matric points and the ability to perform well on the DAT tests. It could be argued that the use of DAT is just reinforcing the matric points as argued by Huysamen (1996; 1999; 2000). However, the opinion of Botha and Cilliers (1999) is also relevant, that we need to generate as much information about the student as possible before making decisions on their academic future. One must also remember that the matric system as it is currently known, will fall away. The fact that the correlation between matric and DAT is good, and between DAT and performance in Biosciences is even better, can be used as an additional entrance criterion in the future[18].

A review of the literature in other countries shows that in the USA males had a slightly higher composite score on the ACT and SAT than females. The most consistent finding is that female students generally attain higher college GPAs than predicted, particularly when the SAT is considered in the admissions process.

In our research, apart from tests 3 and 4 (Non-Verbal Reasoning: Figures and Calculations) and the final results for Chemistry and Physics, females outperformed the males. Females attained a significantly higher average on the DAT tests but the difference between female and male on their final results in Biosciences is not significant (Table 4.4). As evident in Table 4.2 and Figure 4.7 females performed better in the three different combinations of the tests and did particularly well with the combination of test 4,5,6,7[19]. Thus females from this cohort of students are better suited in all these categories. This is a complete reversal, from the data that has been recorded in the USA[20]. Despite this, females appear to be less suited to calculations and Non-Verbal Reasoning. The poor Non-Verbal Reasoning skills of the student impacts directly on the skills necessary for a biologist and thus it is essential that developmental action is taken in this regard.

The Department of Education wishes to increase the participation of blacks and women, especially in the areas of Science, Engineering and Technology. Furthermore, if South Africa is to produce sufficient numbers of graduates with the skills required by a modern economy, the number of postgraduate enrolments will also need to increase. Given these goals for the higher education system, these results are especially encouraging in that it shows that females are eminently suited to study in these fields. This research suggests that the long held belief that males are better suited to these disciplines can be debunked.

---

[18] To substantiate this in Israel the PET scores correlate more highly with college grades than matriculation scores (overall validity 0.38 and 0.32 respectively). However, one must remember that is not the case in the USA.

[19] The first series of tests indicates potential for further study while the middle series of test (4,5,6,7) show commercial clerical aptitude, while the last series shows potential for higher education.

[20] In my experience I have generally found that the female student is more thorough and perseveres to far greater degree.

In their review MacDonald *et al.* (2000) conclude that the nature of the actual tests can be a source of bias, with language being a potential problem. As previously stated language is often regarded as one of the most important moderators with regards to test performance. Poor performance on assessment measures could be more a reflection of language, if the language in which the test is given is not the home language.

Although the students did fairly well in the vocabulary test, when home language was compared to the average of all the DAT tests and the final result for Biosciences (Figure 4.8 and 4.9 and Table 4.6 and 4.7) it was found that language has a profound effect on these two results. Interestingly, students that have English as a home language did not perform the best in the DAT tests nor the Bioscience course. This is probably the result of the Indian students who had an average performance in the DAT series of tests yet have English as their home language. The language designated "?" are a combination of home languages from non South African students. These students performed particularly badly. This is something that must be borne in mind when selecting student from foreign countries. These students should probably have a language test and on the basis of these results attend a compulsory remedial course to improve their English language skills.[21] For the most part, the trends in the language-related DAT tests generally mirror students' performance in the course as a whole. This highlights the difficulties that face the second language student when entering an English medium university.

With respect to English language proficiency, it is well known that South African students typically demonstrate high levels of oral fluency in everyday English but weak reading and writing skills. Wherever possible, we must affirm and build on students' home languages and cultures using relevant examples in the lectures and tasks assigned to the student. Initially, we need to shift gradually from everyday usage to controlling specialised academic discourses. This will require intensive mediation and scaffolding in which students have opportunities to interact with the lecturers and demonstrators in the course.

Although the use of reading and writing courses for those with poor English language skills has been advocated, it is important to note that the learning of languages for academic purposes is labour-intensive and time-consuming. Research shows that 'add-on' remedial courses have little value, particularly because few students are able to transfer what is learnt to the course. Relevance and contextualisation are crucial for effective language learning and recurrent, meaningful and authentic practice is also required. This means that we will have to integrate language proficiency development with curriculum content and this must be developed through real content and contextualised tasks. The design and teaching of this adjusted course will mean collaboration between the lecturers and language development specialists.

---

[21] It must however be noted that some of our foreign student have superb English skills hence the need for the test.

It is very difficult to generalise when it comes to performance (in the DAT) that is linked to language and this needs to be further addressed. When taking ethnic language into account the majority of the students have Zulu as their home language. This group of student did particularly badly in the DAT tests. A great deal has been written about bias as a result of language. However, although language may play a role in affecting test scores, attempts to allow for this have had little success in reducing the score gap McDonald et al. (2000). Thus further research will have to be undertaken so we can factor this into our calculations in future, if we continue to use the DAT battery of tests.

The marked difference in the performances of the different race groups is perhaps the most contentious of all the findings in the research. It is however important to note that this difference cannot be ignored nor overlooked and must be addressed head on. Similar problems occur throughout the world in both first and third world countries. From this research, White students perform significantly better than the other race designations when it comes to the final Biosciences result. Within the categories African, Indian and Coloured there is no significant difference (Table 4.9). When it comes to the different DAT tests, Africans perform significantly worse than the other three groupings and the Indian grouping did significantly worse when compared to Whites. When performing a repeated measure RMANOVA; race, the different tests and the combination of both are significantly different (Table 4.11).

Perhaps in light of the history of the skewed educational provision in this country it may be beneficial to note what has happened in the rest of the world when aptitude tests are used to determine university entrance. Neisser et *al.*, (1996) observed that the lower scores of ethnic minority groups on IQ tests are not unique to America. In other countries which have disadvantaged social groups, similar differences have been observed (e.g. New Zealand, India). Across the world, children who are born into these minority groups tend to do less well in education. Comparisons between ethnic groups show that Whites outperformed all other groups on the verbal section of the SAT. African Americans have the lowest composite score, and also had lower scores on each of the subtests than all other ethnic groups (McDonald *et al.*, 2000). The results from this study showed the same trends. This can probably be explained by the fact that the tests test for developed reasoning abilities in English, the first language of most White test takers. In addition whites in South Africa generally have a more advantaged socio-economic background.

Corley *et al.* (1991) examined SAT differences between students from urban and rural communities and concluded that the differences were not due to differing standards between rural and urban high schools. The same cannot be said for the South Africa situation. There is a clear difference in the quality of education between predominantly Black rural /township schools and affluent White suburban school. This is a legacy of apartheid education; associated with inadequate funding, under resourced schools, poorly qualified teachers in some instances and the difficulties the students must contend with when try to generate a culture of learning. There is also the added burden of not been able to afford school fees, poor nutrition and difficulties in attending schools. As if this was not enough of a burden, these students must also deal with the fact that they must use a second language as the language of learning.

Huysamen (1996; 1999; 2000) and Mitchell and Fridjhon (1987) feel strongly that it would be unfair to apply matric results or use results of aptitude tests for those applicants who had inferior schooling as well as for those of who had privileged schooling. In review McDonald *et al.* (2000) conclude that Blacks and other ethnic groups often score considerably lower on tests such as the SAT and ACT than White students. The evidence on whether these differences are a fair reflection of the subsequent college performance of these ethnic groups is inconclusive.

On reflection the trends in the results that we have obtained are not unique; the trends are similar to other countries, just more exaggerated in South Africa. This result is surprising in light of the social engineering that has happened in the past in our country. There is a large discrepancy in the actual values between the different groupings and this must be further investigated and before any test of this type is implemented, further refining of the norms for the different race groups is essential. To reiterate the sample used to construct the norms tables does not accurately reflect the race distributions in South Africa, and the test constructors admit that there was insufficient representation among black students. Owen *et al.* (2000) states that the norms furnished in the manual are subject to serious shortcomings because of the problems experienced during standardisation of the DAT-L. As regards items of the different tests in general, these functioned best for White learners followed by Indian and Coloured (Owen *et al.*, 2000).

Although there is some doubt about the accuracy of the norms provided for the DAT-L, the data can still be used as initial values in the process of determining an organization's own norms. In the course of time these values can be refined.

The policy in South African higher education is to increase access and admission to Universities and as a result there will be an increasing number of students who will be under-prepared for the University experience. In light of the results from the DAT tests it is obvious that that we must institute a comprehensive response to some of the disturbing results. The following discussion will speculate on a few possibilities that may assist students when entering university.

Firstly, students need to be placed in appropriate learning programmes, be they foundation courses, credit-bearing foundation courses that are structured as part of first year programmes, or augmented degree programmes. Some of the problems that have been highlighted by the research can be addressed through a flexible and responsive curriculum involving ongoing curriculum innovation.

The results from the DAT tests for the ex SFP students was surprising and although they did relatively well in the Biosciences exam their initial DAT scores were poor. Perhaps it would be in the interest of the SFP course to spend more time on the development of the skills that these tests suggest these students lack. As we already have a foundation programme our approach to first semester could serve as an intentional bridge between school and university studies. There should be scaffolding with additional academic support and mentoring. It is important that we create a learning environment which focuses on learner activity, interaction, variation in task demands and the application of knowledge to the real world. This should facilitate the development of these high order cognitive skills in which the findings suggest students are generally weak.

Our course must cater for the learning needs of its target student population, ensuring that all students have a fair chance of success. The design of the course and delivery methods must be varied and flexible enough to accommodate a diversity of students.

Good teaching is open to change; it involves constantly trying to find out what the effects of teaching are on learning, and modifying that teaching accordingly. This includes: an understanding what a student will be expected to have upon arrival; key skills; cognitive skills; subject specific skills; skills that are socially situated; and whether there are many sets of basic skills available and operational. This involves a practical approach and possibly the creation of specific skills development modules. This is especially relevant for Non-Verbal Reasoning and Reading Comprehension which are particularly weak in South African students

Practical approaches that we could use include; one-on-one teaching and maximization of interactions; active learning; introductory courses; and a "Biology course for non-majors". With the limitations of our student body it is important that we give up trying to cover large content areas and rather concentrate on teaching the skills and thinking processes that enable students to acquire and synthesize knowledge.

Although the ideal is to increase access, this is clearly a labour-intensive undertaking and so we must be acutely aware of the constraints placed upon us; any curriculum changes must take into account the limitations of institutional capacity in terms of physical, financial and human resources.

In the wider South African context higher education needs to benchmark the FETC against the critical competencies required of students at university. This could be achieved by developing benchmark tests to test for academic literacy, numeracy and mathematical competencies that are critical for success in Higher Education studies. In this regard, the DAT battery of tests does show promise, especially for biological subjects. However it is necessary for further research to be conducted so that these tests are shown to be scientifically valid and reliable.

In summary the findings of this research project suggest the following:

- Do the DAT-L tests have predictive validity for academic success in first year biology in South Africa? To a large extent they do.
- Can these tests be used to predict successful academic study of Biology? Based on the preliminary results from this research, yes they can.
- Can the DAT tests be used to complement school leaving results in making judgments about selection and placement? Initial results suggest they there is still a great deal more research needed, especially in terms of the norms tables that have been published, before the DAT-L tests can be used as benchmark tests for entry into higher education in South Africa.

6 Concluding observations

Results from the research on the DAT-L battery of tests have highlighted that some of our students have a deficit when it comes to certain skills necessary for studies in the Biological sciences. For example the students performed particularly badly in the Non-Verbal Reasoning and Comparison tests. One of the features the tests did show was that English language proficiency plays a large role in succeeding in these tests. It takes longer for students that have English as a second language to process information. As a result poor performance could be a reflection of problems associated with language difficulties and not weak individual ability. This means that we will have to integrate English language proficiency development with curriculum content and this must be developed through real content and contextualised tasks. The design and teaching of this adjusted course will mean collaboration between the discipline lecturers and language development specialists.

Language can easily become a source of test bias. This highlights the difficulties that face the second language student when entering an English medium university. With respect to English language proficiency, it is well known that South African students typically demonstrate high levels of oral fluency in everyday English but weak reading and writing skills. The Reading Comprehension test aims to measure the learner's ability to comprehend what he /she is reading. The choice of passages in the DAT-L test is perhaps too advanced for some of the second-language first year students and the contexts of the readings are too Eurocentric. In addition, the time available to complete the test may also have contributed to students' poor performance. Thus it is probable that this test is biased against second-language students.

Females attained a significantly higher average on the DAT-L tests but the difference between female and male on their final results in Biosciences is not significant. White students performed significantly better than the other race designations when it came to the final Biosciences result. When it came to the different DAT tests, Africans performed significantly worse than the other three groupings and the Indian grouping did significantly worse when compared to Whites. However, these results are distorted because of the English proficiency backlog of some of our students.

When comparing the matric points with the DAT and the final mark, there was a better correlation with the DAT tests than with the final mark. What this highlights is the fact that in this research, the use of the results from the aptitude tests gives a better indication (predictor) of the performance of the student in Biosciences than if the matric points were used to predict performance.

The trend where biology shows a good correlation with high school marks and aptitude tests appears to be world wide. A major problem however, is that if we were to use the DAT as a means of selecting students it would be effective for biology but not as effective for the other basic science subjects.

Despite the good correlation with biology, it must be accepted that the study has a number of limitations. The correlations reported in the study are for the Biosciences and the sample size is fairly small. Although a number of pilot studies have been conducted, this data was not included. To improve the reliability of the results, the results from a number of different student cohorts should be included. To make more meaningful correlations with the other subjects, it would be necessary to give the test to the whole of the first year Science and Agriculture intake rather than limiting it to those students doing these subjects in conjunction with biology.

Aptitude is impossible to measure directly since, at best; aptitude tests can measure an individual's abilities at a given moment. In considering an aptitude test, it is never a simple matter to declare it valid or not valid. It is also necessary to look at the strength of its predictive validity, one aspect of validity, and to consider whether its predictive ability can be generalized across a range of situations.

In order for test scores to be used for guidance and placement, it is necessary that they have some predictive validity. In countries which have disadvantaged social groups, children who are born into these groups tend to do less well in education and to drop out earlier. In this study, whites tended to outperform all other ethnic groups. African students performed very poorly on the tests, having the lowest average scores of any ethnic group. Socio-economic status is likely to be a contributory factor to score differences, particularly in a country of great social inequality such as South Africa. There is a clear difference in the quality of education between predominantly Black rural /township schools and affluent White suburban schools. This is a legacy of apartheid education; associated with inadequate funding, under resourced schools, poorly qualified teachers in many instances and the difficulties the students must contend with when trying to generate a culture of learning. As if this was not enough of a burden, these students must also deal with the fact that they must use a second language as the language of learning. Variables such as these interfere with the predictive validity of an aptitude test and are very difficult to accommodate in the design of the test itself.

However, the major limitation of this study was that the original sample used to construct the norms tables for the DAT-L tests, did not accurately reflect the race distributions in South Africa. Owen *et al.* (2000) admit that there was not sufficient representation among black students across the educational spectrum. The black student samples were skewed as only rural 'self-governing' schools were used as participants and this sample is not representative. In fact, Owen *et al.* (2000) state that the norms furnished in the manual are subject to serious shortcomings because of the problems experienced during standardisation of the DAT-L, these included unrest at many black schools and many matriculants were preparing for the final examination. Because of these problems, the norms should be regarded as a broad guideline only.

These limitations of the study suggest that further test development and research will be necessary before we have a sufficiently valid, reliable and fair instrument that can be used to test and place students entering higher education from across the spectrum of educational and socio-economic backgrounds that exist in South African society.

Kilfoil (1999) and Botha and Cilliers (1999) confirm the unreliability of school results and stress that the assessment of Higher Education applicants is becoming increasingly important. Aptitude tests to assist in the selection and placement of students for University study, does not provide a perfect solution to this problem, but with further refinement, they could provide one instrument that could be used to address the problem of selection.

References

American College Testing Program (2000). Selections from the 2000 National Score Report: Academic Abilities and Nonacademic Characteristics of ACT-Tested Graduates (2000 ACT National and State Scores). Available: http://www.act.org/news/data/00/00data.html [October, 2004]

Bailey, K.D. (1987). Methods of Social Research. 3rd ed. New York, The Free Press

Barnard, S. (1992). Multilingualism. South African Journal of Higher Education, 6 (3): 5

Behr, A.L. (1985). The senior certificate examination as a predictor of university success. Journal for education, 5 (3): 107

Beller, M. (1994). 'Psychometric and social issues in admissions to Israeli universities', Educational Measurement: Issues and Practice, 13 (2): 12

Bokhorst, F.D., Foster, D.H. and Lea, S.J. (1992). Factors affecting academic performance in first-year psychology at the University of Cape Town. South African journal of higher education, 6 (1): 59

Botha, H.L. and Cilliers, C.D. (1999). Preparedness for university study: designing a thinking skill test. South African Journal Of Higher Education, 13 (1): 144

Braun, H., Ragosta, M. and Kaplan, B. (1986). The Predictive Validity of the Scholastic Aptitude Test for Disabled Students (ETS Research Report 86 - 38). Princeton, NJ: Educational Testing Service. Cited by: McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2000). Aptitude Testing for University Entrance: A Literature Review. Available: http://www. suttontrust. com/text/SAT-Literature% 20Report.pdf [August 2004]

Bridgeman, B. and Wendler, C. (1991). 'Gender differences in predictors of college mathematics performance and in college mathematics course grades. Journal of Educational Psychology, 83 (2): 275

Bridgeman, B., McCamley-Jenkins, L. and Ervin, N. (2000). Predictions of Freshman Grade-Point Average from the Revised and Recentered SAT I: Reasoning Test (College Board Research Report No.2000 - 1/ETS Research Report 00 - 1). New York, NY: College Board. Cited by McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2000). Aptitude Testing for University Entrance: A Literature Review. Available: http://www.suttontrust.com/text/SAT-Literature%20Report.pdf [August 2004]

Brown, C.R. (1995). The Effective Teaching of Biology. London, Longman

Chapman, K. (1996). Entry qualifications, degree results and value-added in UK universities. Oxford Review of Education, 22 (3): 251

Chua, W.F. (1986). Radical Developments in Accounting Thought. The Accounting Review. 61: 601

Corley, E.R., Goodjoin, R. and York, S. (1991). Differences in grades and SAT scores among minority college students from urban and rural environments. The High School Journal, 74 (3): 173

Crouse, J. and Trusheim, D. (1988). The Case Against the SAT. Chicago, The University of Chicago Press

Custer, R.L. (1996). Qualitative Research Methodologies. Journal of Industrial Teacher Education 34(2): 3

DeVilliers, L. (1977). Design and Standardization of Scholastic Aptitude Tests for Indian South Africans. Institute for Psychometric Research. Pretoria: Human Sciences Research Council

Fierro, J. (2004). Quantitative research. Available: http://www.gsu.edu/~mstswh/courses/it7000/papers/quantita.htm [September 2004]

Fincher, C. (1990). Trends in the Predictive Validity of the Scholastic Aptitude Test (ETS Research Report 90 - 13). Princeton, NJ: Educational Testing Service. Cited by: McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2000). Aptitude Testing for University Entrance: A Literature Review. Available: http://www. Suttontrust .com/text/SAT-Literature%20Report.pdf [August 2004]

Fleming, J. and Garcia, N. (1998). Are standardized tests fair to African Americans? Predictive validity of the SAT in black and white institutions. The Journal of Higher Education, 69 (5): 471

Foxcroft, C. (2004). The importance of benchmarking the FETC for the Higher Education sector. Izwi. Voice of HE Leadership. South African Universities Vice Chancellors Association, 2: 9

Griesel, H. (2001). Institutional views and practices: Access, admissions, selection and placement. A summary of findings. In: The Challenges of Access and Admission. South African Universities Vice Chancellors Association, p 17

Harris, A.M. and Carlton, S.T. (1993). Patterns of gender differences on mathematics items on the Scholastic Aptitude Test. Applied Measurement in Education, 6 (2): 137

Hopkins, W.G. (2000). Quantitative research design. Available: http//: www. sportsci.org/jour/0001/wghdesign.html [September 2004]

Hull, C.L. (1926). Aptitude Testing. London: George Harrap and Co.

Huysamen, G.K. (1996). Amerikaanse keuringsmodelle: is dit toepaslik? Bulletin, 3 (1): 12

Huysamen, G.K. (1999). Psychometric explanations for the poor predictability of the tertiary-academic performance of educationally disadvantaged students. South African Journal Of Education, 13 (1): 132

Huysamen. G.K. (2000). The differential validity of matriculation and university performance as predictors of post-first-years' performance. South African Journal Of Education, 14 (1): 146

Hyde, J.S. and Linn, M.C. (1988). Gender differences in verbal ability: a meta-anlysis. Psychological Bulletin, 104 (1): 53

Jencks, C. (1998). Racial bias in testing. In: JENCKS, C. and PHILLIPS, M. (Eds) The Black-White Test Score Gap. Washington, DC: The Brooking Institution.

Jordaan, J.J. De V. (1990). Attitude To Studies Of First-Year University Students In Biological Sciences At The Commencement Of University Level Tuition University Of Cape Town. M.Ed. University of Cape Town

Kilfoil, W.R. (1999). The linguistic competence of science students. South African Journal Of Higher Education, 13 (1): 47

Kotecha, P. (2001). Access and flexibility: Challenges for FET and HET. In: The Challenges of Access and Admission. South African Universities Vice Chancellors Association, p 11

Lowe, H. and Cook, A. (2003). Mind the Gap: are students prepared for higher education? Available: http:// www.swetswise.co./ eAcess/viewTitelssues. do?titleID=112676 Journal of Further and Higher Education. 27(1) 53 [August 2004]

Luckett, K. (2004 pers. com). School of Education Training and Development. University of KwaZulu-Natal.

Macklem, G.L. (1990). Measuring Aptitude: Practical Assessment, Research and Evaluation. American Institute for Research. Eric Clearinghouse. Available: http://www.ericae.net/pare/getvn.asp [September 2004]

McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2000). Aptitude Testing for University Entrance: A Literature Review. Available: http://www.suttontrust.com/text/SAT-Literature%20Report.pdf [August 2004]

Mellanby, J., Martin, M. and O'Doherty, J. (2000). The "gender gap" in final examination results at Oxford University. British Journal of Psychology, 91(3): 377

Miller, R. (1992). Double, double, toil and trouble: the problem of student selection. South African Journal of Higher Education, 6 (1): 98

Mitchell G. and Fridjhon P. (1987). Matriculation and university performance. South African Journal Of Science, 83: 555

Morgan, R. (1990). Predictive Validity within Categorizations of College Students: 1978, 1981, and 1985 (ETS Research Report 90 - 14). Princeton, NJ: Educational Testing Service. Cited by: McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2000). Aptitude Testing for University Entrance: A Literature Review. Available: http://www.suttontrust.com/text/SAT-Literature%20Report.pdf [August 2004]

Mouton, J. (1996). Understanding social research. Van Schaik. Pretoria

Myers, M. D. (1997). Qualitative Research in Information Systems. MIS Quarterly. 21(2): 241

Naude, C. (2003). Waiting for the new matric. The Natal Witness Publish Date: 11 August 2003

Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J. and Urbina, S. (1996). Intelligence: knowns and unknowns. American Psychologist, 51(2): 77

Nisbet, B. (2000). Psychometric Tests. University of Sussex: Career Development Unit. Available: http://www.sussex.ac.uk /unts/CDU/ psych.html. [September 2004]

Orlikowski, W.J. and Baroudi, J.J. (1991). Studying Information Technology in Organizations: Research Approaches and Assumptions. Information Systems Research. 2: 1

Owen, K. and Vosloo, H.N. (1999). Senior Aanlegtoetse Vorm L, Senior Aptitude Tests Form L. Pretoria, Human Sciences Research Council

Owen, K., Vosloo, H.N., Claassen, N.C.W. and Coetzee, H. (2000). Manual for Differential aptitude tests Form L (DAT-L) = Aanlegtoetse Vorm L (DAT-L). Pretoria, Human Sciences Research Council

Phala, J.M. (1992). Input quality control in higher education. South African Journal Of Higher Education, 6 (3): 73

Ragosta, M., Braun, H. and Kaplan, B. (1991). Performance and Persistence: a Validity Study of the SAT for Students with Disabilities (College Board Report No.91 - 3). New York, NY: College Board. Cited by: McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2000). Aptitude Testing for University Entrance: A Literature Review. Available: http://www. suttontrust.com/text/SAT-Literature%20Report.pdf [August 2004]

Rosser, P. (1989). The SAT Gender Gap: Identifying the Causes. Washington, DC: Center for Women Policy Studies. Cited by: McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2000). Aptitude Testing for University Entrance: A Literature Review. Available: http://www.suttontrust.com/text/SAT-Literature%20Report.pdf [August 2004]

Roznowski, M. and Reith, J. (1999). Examining the measurement quality of tests containing differentially functioning items: do biased items result in poor measurement? Educational and Psychological Measurement, 59 (2): 248

SAUVCA (2003). The proposed National Curriculum Statements Grades 10-12 (Schools). Draft University Sector Response. South African Universities Vice Chancellors Association

Schaffner, P.E. (1985). 'Competitive admission practices: when the SAT is optional', Journal of Higher Education, 56 (1): 55

Schmitt, A.P. and Dorans, N.J. (1990). 'Differential item functioning for minority examinees on the SAT', Journal of Educational Measurement, 27 (1): 67

Schneider, D. and Dorans, N. (1999). Concordance Between SAT I and ACT Scores for Individual Students (Research Notes 07). New York, NY: College Board. Cited by: McDonald, A.S., Paul E. Newton, P.E., Chris Whetton, C. and Benefield, P. (2000) Aptitude Testing for University Entrance: A Literature Review. Available: http://www.suttontrust.com/text/SAT-Literature% 20Report.pdf [August 2004]

Schurr, K.T., Henriksen, L.W. and Ruble, V.E. (1990). 'The use of the College Board classification of high schools in predicting college freshman grades', Educational and Psychological Measurement, 50 (1): 219

Shayer, M. and Adey, P. (1981). Towards a Science of Science Teaching. London, Heinemann Educational Books

Stricker, L.J., Rock, D.A. and Burton, N.W. (1996). Using the SAT and high school record in academic guidance. Educational and Psychological Measurement, 56 (4): 626

Super, D.E. (1949). Appraising Vocational Fitness. USA: Harper and Brothers.

Swanepoel, S. (2003). The predictive validity of the Myers-Briggs for academic success of students. Proceedings of the: International Conference on Social Sciences, Hawaii

Van Wyk, J.A., and Crawford, J.L. (1984). Correlation between matric symbols and marks obtained in a first-year ancillary physics course at the University of the Witwatersrand. South African Journal Of Science, 80: 8

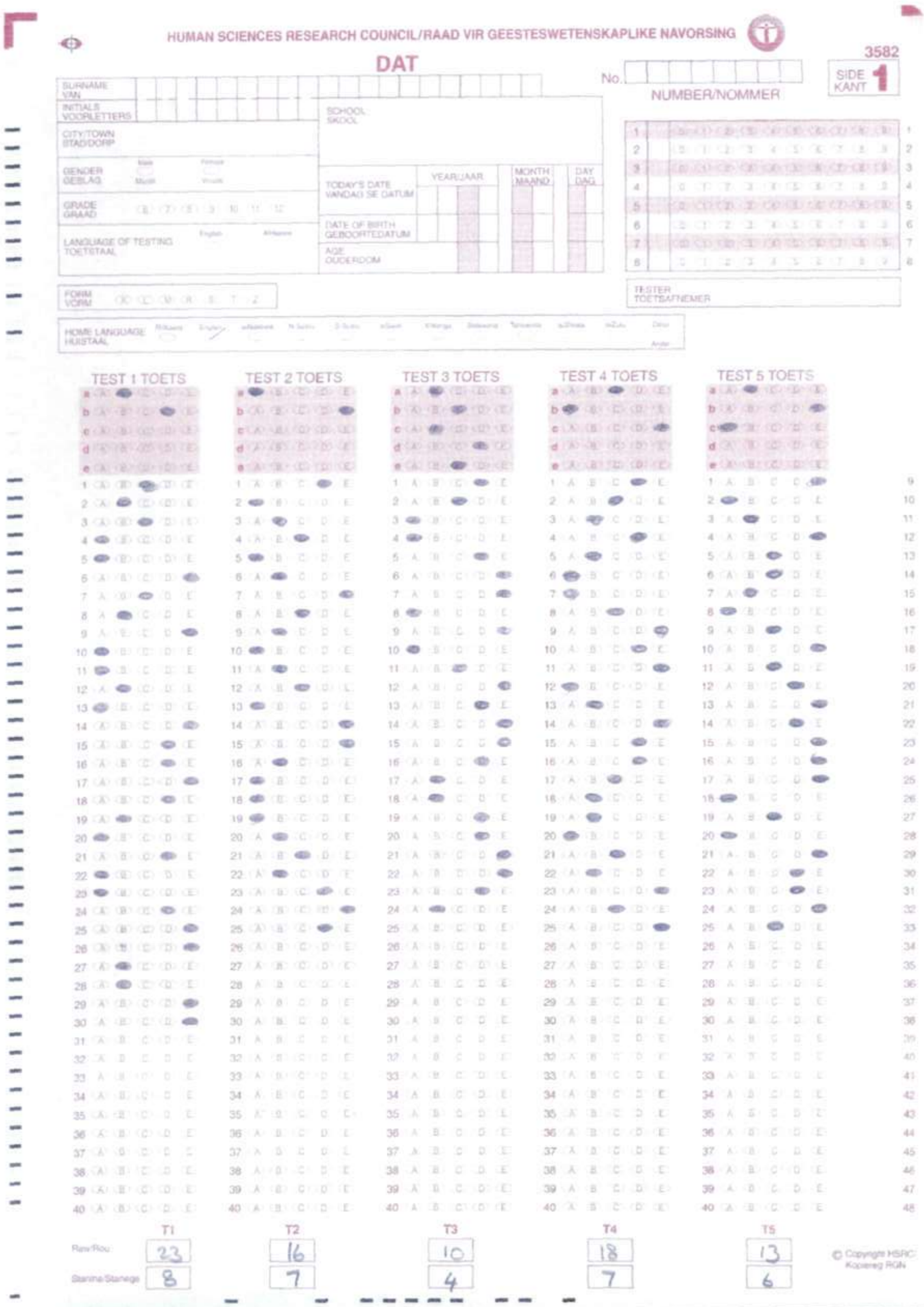
Williams, E.M. (1950). An investigation of the value of Higher School Certificate results in predicting performance in first-year university examinations. British Journal of Educational Psychology, 20: 83.

Wilmouth, D. (1991). Should the SAT be a Factor in College Admissions? (ED 345 592). Washington, DC: U.S. Department of Education, Educational Resources Information Centre. Cited by: McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2000). Aptitude Testing for University Entrance: A Literature Review. Available: http://www.suttontrust.com/text/SAT-Literature%20Report.pdf [August 2004]

Zeidner, M. (1987). Age bias in the predictive validity of scholastic aptitude tests: some Israeli data. Educational and Psychological Measurement, 47(4): 1037

Appendix 1

Typical test sheet as filled in by the students (Side 1 of 2) and scored.


HUMAN SCIENCES RESEARCH COUNCIL/RAAD VIR GEESTESWETENSKAPLIKE NAVORSING

DAT

3582

SIDE KANT 1

No. NUMBER/NOMMER

SURNAME VAN

INITIALS VOORLETTERS

SCHOOL SKOOL

CITY/TOWN STAD/DORP

GENDER GESLAG

GRADE GRAAD

LANGUAGE OF TESTING TOETSTAAL

TODAY'S DATE VANDAG SE DATUM

YEAR/JAAR MONTH MAAND DAY DAG

DATE OF BIRTH GEBOORTEDATUM

AGE OUDERDOM

TESTER TOETSAFNEMER

FORM VORM

HOME LANGUAGE HUISTAAL

TEST 1 TOETS | TEST 2 TOETS | TEST 3 TOETS | TEST 4 TOETS | TEST 5 TOETS

	T1	T2	T3	T4	T5
Raw/Rou	23	16	10	18	13
Stanine/Stanege	8	7	4	7	6

Appendix 2. Entrance Requirements (Taken from University of KwaZulu-Natal Website)

For South African School Leaving Candidates

Early Selection

Applicants with outstanding Grade 11 (Std 9), Matric June/Trial on mock results may be selected on submission of a certified copy of results. Selections will be final provided selected applicants obtain a matriculation exemption. Other applications will be considered after the release of the matriculation results.

Selection on Matriculation Point Score

Selection commences when matriculation results are released and continues until all available places are allocated.
Students are admitted on the basis of academic merit, which is assessed by converting Senior Certificate symbols into a point score.

Once the examination results of all applicants are known, scores are calculated and ranked in descending order of scores for their choice of programme. Selection takes place from the top down, until the number of applicants who may be accepted for that programme is reached. The score at which this is reached is known as the 'cut-off point', and this varies from programme to programme.

Applicants below that point score are normally not selected. This total may change from year to year, depending on the number and quality of applications.

Applicants to MBChB should note the different method for calculation of point scores for MBChB. more>

The University will receive Senior Certificate/matriculation examination results from the Central Applications Office. It is the responsibility of all applicants to ensure their preliminary statements of symbols reach the Central Applications Office as soon as possible. Applicants must write their degree choice and CAO reference number on the statement. Applicants who are offered places at the University will receive letters from the University advising them of our offer. Applicants will be required to pay an acceptance deposit if they want to secure their places (R500 for 2004).

Faculty Entrance Requirements

In addition to the general entrance requirements, some Faculties have laid down specific entrance requirements. These are listed under Faculty Information in this publication. Also refer to the Entrance Requirements tables for individual faculties.

Submission of Senior Certificate

It is important to note that you must submit a copy of your Senior Certificate to your Faculty Office on or before 31 May 2005.

2004 Matriculants who qualify for Ordinary Conditional Exemption must contact their school or apply direct to the Department of Education for a Conditional Exemption Certificate. These certificates will stipulate the conditions which have to be met to enable them to qualify for a full exemption. Applicants who need to apply for foreign or mature age exemption must contact their Faculty Office after registration for assistance.

Calculation of Matriculation Point Score

The point score is calculated by matching each of the six best Matriculation subject symbols with the relevant value listed either under higher grade (HG) or standard grade (SG) and then adding the scores to give the total. Full credit is given for Additional Mathematics, but Advanced Mathematics is not considered. Functional Mathematics and Functional Physical Science offered on Standard Grade will be considered as SG subjects for the calculation of your point score, but will not satisfy entrance requirements for programmes with a required pass in Mathematics or Physical Science at standard or higher grade.

If a seventh subject was passed with a symbol of at least 'E' on higher grade or 'D' on standard grade, a bonus of two points will be added to the score. If an 8th subject has been passed it will strengthen the application. It is important to note that the bonus points will not apply for applications to Medicine.

If a subject written on higher grade has been given a condoned standard grade pass, the points score is to be calculated using the standard grade value.

Points Table		
Symbol	HG	SG
A	8	6
B	7	5
C	6	4
D	5	3
E	4	2
F	3	1

Example			
Subject	Grade	Symbol	Points
English	HG	B	7
Afrikaans	HG	C	6
Mathematics	HG	A	8
Physical Science	HG	D	5
History	SG	C	4
Geography	HG	E	4
TOTAL			34

General Entrance Requirements

The basic legal requirement for admission to degree studies at University in South Africa is a matriculation exemption. Applicants need to be in possession of one of the following:

1. A Senior Certificate with matriculation exemption endorsement.

2. A Certificate of Conditional Exemption. Matriculants who did not obtain a full exemption and who wish to apply for a Conditional Exemption Certificate must contact their school and the Department of Education.

Applicants must contact the Admissions Officer to establish if the programme chosen will consider applicants with conditional exemption.

3. Applicants with a Senior Certificate who do not qualify for a Conditional Exemption Certificate may apply for admission through an alternative access programme, and on successful completion of the programme the University will apply for a Conditional Senate Exemption Certificate.

4. Mature Age Exemption: Applicants over the age of 23 may qualify for a conditional exemption certificate on grounds of mature age. Certain subject requirements apply and applicants must contact their Faculty Offices for more information. The University will apply to the Matriculation Board for an exemption certificate after the applicant has registered at the University.

5. Applicants in possession of non-South African school leaving certificates must enclose certified copies of documents and translations where necessary. On application the University will contact the Matriculation Board to determine if applicants qualify for Foreign Conditional Exemption Certificates and assist students in applying for an exemption certificate after they have registered.

6. For admission to certain programmes the possession of a Senior Certificate (without exemption) is sufficient.