

https://dspace.org/

The predictive validity of the mental alertness, reading comprehension, arithmetic reasoning and conceptual reasoning tests as used by the Wits Business School.

Maduma, Eunice Sibongile Sylvia

2012-06-29
http://hdl.handle.net/10539/11566

Downloaded from a DSpace repository. https://dspace.org/

THE PREDICTIVE VALIDITY OF THE MENTAL ALERTNESS, READING COMPREHENSION, ARITHMETIC REASONING AND CONCEPTUAL REASONING TESTS AS USED BY THE WITS BUSINESS SCHOOL

Eunice Sibongile Sylvia Maduma

9700877K

A Research Report submitted to the Faculty of Humanities, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Arts in Psychology by Coursework and Research Report. December, 2011

Declaration

I hereby declare that this dissertation is my own work and has not been submitted to any institution in fulfilment of another degree.

Signature: _____ Date: 22 December 2011

Acknowledgements

Several people have made significant contributions towards the successful completion of this study. I wish to extend grateful acknowledgement to them:

- My supervisor, Professor Charles Potter: Words cannot express how grateful I am for your patience and guidance throughout the years.

- Professor Peter Fridjhon: Knowing my limitations with statistics, you were willing and available to make sure that I knew what I was doing as I administered and interpreted the statistics.

- Staff at the WBS Faculty Office and Academic Programmes Office, more so Mrs Zimasa Koyana for responding to my endless questions.

- Yet the determination and motivation to complete my studies would not have existed if it were not for the encouragement from my family, most importantly my husband, Matije. I would not have achieved this without your support.

- Lastly, my children Tswelopele and Kitso for understanding when mommy had to work late and over weekends to finalise the write up of the report.

Love....

Abstract

In this research, the scores on four National Institute for Personnel Research (NIPR) tests (the Mental Alertness, Reading Comprehension, Arithmetic Reasoning and Conceptual Reasoning Tests) formed the independent variables and course results on the Master of Business Administration (MBA) programme the dependent variables. The analysis involved examining data of 741 MBA students to determine the relationship between the four NIPR tests and academic performance in the ten core courses of the MBA as one dependent variable, and academic success in completing the MBA degree as a second dependent variable. The language background and racial classification of the students were also examined in relation to the two dependent variables of academic performance and academic success.

The results from stepwise regression indicated that Mental Alertness was the best predictor of academic performance overall but that this explained only 10% of the variance in academic performance. Adding the additional predictor variables into the regression equation added only an additional 2% to the amount of variance which could be predicted through use of the four NIPR tests in combination. Arithmetic Reasoning Test was the only test that was not accepted into the regression model. Logistic regression showed that none of the variables were significant in predicting whether students would successfully complete their MBA studies or fail to. None of the four NIPR tests were able to meet the significance level necessary to be entered into the model.

A number of differences in academic performance were found which could be attributed to language preference. It was evident that the group that preferred using English for communication had higher odd ratios than the group that did not prefer English as a language of communication. However, the effect sizes of these differences were small.

With regards to differences between the academic performances of the different racial groups, it was evident from the analyses that Black students were outperformed by White and Asian students. The difference between Asian and White mean scores was very small compared to the difference between mean scores of Black and White students, and the effects were medium to large confirming significant differences between the three racial groups. In terms of differences in the academic success of different racial groups, White students had better odds of successfully completing their studies than Black students. In addition, Asian students also had better odds of succeeding than Black students.

The overall conclusion from this study was that, while the abilities tapped by all four NIPR tests contributed towards determining academic performance, their overall contributions were not high. They could only account for 12% of the variation in academic performance of the Wits MBA student sample. In contrast, extraneous variables explained 88% of the variance in academic performance. This indicated that the four NIPR tests contributed information which was only marginally useful to the selection decision.

Table of Contents

Declaration....1
Acknowledgements....2
Abstract....3
Table of Contents....5
1. Introduction....9
1.1. Overview....9
1.2. Focuses of this Study....10
1.3. Aims and Rationale of this Study....11
1.4. Research Questions....13
1.5. Summary....14
2. Background and Context of the study....15
2.1. The Context of the Study....15
2.2. What is a Master of Business Administration?....16
2.3. MBA Courses....17
2.4. Costs of Studying for an MBA....18
2.5. Implications for the Current Study....20
2.6. Summary....21
3. Literature Review....24
3.1. Introduction....24
3.2. The Importance of the Employment Equity Act in Prediction Studies....26
3.3. The Cognitive and Non-Cognitive Variables as Predictors of Academic Success....28
3.3.1. Non Cognitive Variables....28
3.3.2. Cognitive Variables....29
3.4. The Role of Language in Predicting Academic Success....33
3.5. The Role of Race in Predicting Academic Success....35
3.6. Theoretical Perspectives into Intelligence Testing....36
3.6.1. Static Assessment....39
3.6.2. Dynamic Assessment....40
3.7. Summary....42
4. The National Institute of Personnel Research (NIPR) Group Tests....44
4.1. Background....44
4.2. The NIPR High Level Battery (B/75)....45
4.2.1. Mental Alertness Test....46
4.2.2. Reading Comprehension Test....47
4.2.3. Arithmetic Reasoning Test (A/131)....49
4.2.4. Conceptual Reasoning Test (1/0/1990)....51
4.3. South African Research on the Above NIPR Tests....53
4.3.1. Reliability of the Four NIPR Tests....53
4.3.2. Literature on the Use of the Four NIPR Tests in Combination....54
4.3.3. Literature on the Use of the High Level Battery Mental Alertness Test....56
4.3.4. Literature on the Use of the Arithmetic Reasoning Test....58
4.3.5. Literature on the Use of the Conceptual Reasoning Test....59
4.4. Summary....60
5. Research Methodology....61
5.1. Introduction....61
5.2. Research Questions and Hypotheses....62
5.3. Assumptions....63

5.4. Research Design 64
5.5. An Evaluative Focus 65
5.5.1. A Non-Experimental Research Design Focused on Available (Archival) Data 66
5.6. Procedure for Data Collection 67
5.7. Sampling Methodology 68
5.7.1. Characteristics of the Sample 70
5.8. Procedure for Data Analysis 71
5.8.1. Descriptive Statistics 71
5.8.2. Summary of Descriptive Statistics 75
5.8.3. Inferential Statistics 78
5.9. Instruments 87
5.9.1. Descriptive Data on the Four NIPR Tests 89
5.10. Overall Trends in the Distributions of Scores on the Independent Variables 93
5.11. Academic Performance and Academic Success 93
5.12. Procedure 95
5.13. Summary 96
6. Results and Interpretation 99
6.1. Introduction 99
6.2. Assumptions of Parametric Measures 99
6.3. Results Relating to Research Question 1 102
6.3.1. Relationships between the NIPR Tests and Academic Performance 103
6.3.2. Relationships Between the NIPR Tests and the Ten Core Courses 105
6.4. Results Relating to Research Question 2 108
6.4.1. Assumptions of Multiple Regression 109
6.4.2. Assumptions of Logistic Regression 120
6.5. Results Relating to Research Question 3 124
6.5.1. Results from Stepwise Multiple Regression with Language and Race as Classification Variables 125
6.5.2. Discussion 132
6.5.3. Results from Logistic Regression with Language and Race as Classification Variables 134
6.5.4. Results from Post Hoc Testing of Differences within the Language 138
6.5.5. Results from Post Hoc Testing of Differences within the Racial Classification Variable 140
6.6. Conclusion 145
7. Conclusions, Limitations and Recommendations for Future Studies 147
7.1. Introduction 147
7.2. Summary of Key Findings in the Study 148
7.3. Strengths of the Study 151
7.4. Limitations of the Study 151
7.5. Suggestions for Future Research 155
7.6. Conclusion 155

Reference List 157
Appendices 172

APPENDIX A: Description of Principle and Core Courses
APPENDIX B: Normal Distribution Diagrams

APPENDIX C: Box Whisker Plots
APPENDIX D: Correlation Matrix
APPENDIX E: Scatterplots of Test Scores
APPENDIX F: Results from Stepwise Multiple Regression
APPENDIX F1: Results from Stepwise Multiple Regression Using the Language Variable
APPENDIX F2: Results from Stepwise Multiple Regression Using the Race Variable
APPENDIX G: Results from Stepwise Logistic Regression
APPENDIX H: Results from Two Independent Sample T-Test Analysis
APPENDIX I: Results from Analysis of Variance (ANOVA)

List of tables

Table 1: Reliability Results for the Four NIPR Tests.....54
Table 2: Research Questions, Hypotheses and Associated Statistics.....62
Table 3: Descriptive Statistics of the Demographic Variables.....71
Table 4: Racial Distributions of those who were Successful versus those who were not Successful to Complete their MBA studies.....75
Table 5: Summary Statistics of the Scores for the Independent Variables.....76
Table 6: Results from Kolmogorov Test of Normality.....101
Table 7: Results of Relationships Between NIPR Tests and the Average of the Core Courses.....103
Table 8: Results of Relationships Between the NIPR Tests and the Individual MBA Core Courses.....105
Table 9: Table of Results from Collinearity Diagnostics.....113
Table 10: Results from Multiple Regression Analysis Between NIPR Tests and Academic Performance.....117
Table 11: Results from Forward Stepwise Multiple Regression Analysis 118
Table 12: Results from Logistic Regression Analysis.....123
Table 13: Test Means of the Successful vs. Non-Successful Students with Reference to the NIPR Tests.....124
Table 14: Results from Forward Stepwise Regression for English Speaking Students.....127
Table 15: Results from Forward Stepwise Regression for Non-English Speaking Students.....128
Table 16: Results from Forward Stepwise Regression for Asian Students.....129
Table 17: Results from Forward Stepwise Regression for Black Students.....130
Table 18: Results from Forward Stepwise Regression for White Students.....131
Table 19: Odd Ratios Table for Language Groups 135
Table 20 Odd Ratios Table for Racial Groups.....136
Table 21: Results from *t*-tests.....138
Table 22: Details of Effect Sizes for the Two Language Groups.....139
Table 23: NIPR Tests and Academic Performance Distributions by Race.....141
Table 24: Results from ANOVA Analysis.....142
Table 25: Details of Effect Sizes for the Three Racial Groups.....144

List of Figures

Figure 1: Maslow's Hierarchy of Needs.....16
Figure 2: Distribution of Qualifications of WBS MBA Students.....74
Figure 3: Normal Distribution Graph for the Mental Alertness Test.....89
Figure 4: Normal Distribution Graph for the Reading Comprehension Test.....90
Figure 5: Normal Distribution Graph for the Arithmetic Reasoning Test.....91
Figure 6: Normal Distribution Graph for the Conceptual Reasoning Test.....92
Figure 7: Normal Distribution Graph for the Scores of the Average of Core Courses.....94
Figure 8: Distribution of Residuals for the Average of Core Courses.....110
Figure 9: Results of Influence Diagnostics for the Average of Core Courses.....111
Figure 10: Results of Influence Diagnostics for the Average of Core Courses.....112
Figure 11: Residuals by Predicted for Average of Core Courses.....115

1. Introduction

1.1. Overview

The use of psychometric tests in the selection of students into institutions of higher education is on the increase (Nzimande, 1995). A number of studies, (for example, Buntin, 1994; Holburn, 1991b; MacFarlane, 2006; Nzimande, 1995) have indicated that the majority of these tests and the decisions taken are influenced by the social and political dispensations of a specific era. In South Africa, the apartheid era influenced the decisions that were taken around access and distribution of resources where education was concerned. MacFarlane (2006) referred to the Bantu education system (the apartheid education system for Black people) and how it enforced educating Black township students in Afrikaans and made sure that they had limited or no interaction with English speaking students. Hence it was argued that the system made it difficult for Black students to access "White" institutions.

Potter's study (1991) conducted at the University of the Witwatersrand for example, examined the academic performance of White and Black students in the university's Engineering Faculty. Legislation at the time required Black students to spend at least a year at a segregated Black university before entering Wits. Once at Wits, the Group Areas Act was then applied preventing White and Black students from staying in the same university residences in the same group area. This effectively prevented Black students from accessing the informal study networks which operated out of the White university residences (Potter, Meyer, Scott & Da Silva, 1984).

Consequently "changes in the higher education in South Africa led to the broadened access and the resulting diversity of learner population" (Koch, Foxcroft, & Watson, 2001, p. 126). The

advent of democracy in 1994 brought in a new dispensation and new legislation which enabled a move away from an elitist approach which excluded the majority of students, to mass opportunity where all students stood an equal chance of being accepted (Killen, Marais, & Loedoff, 2003).

The implications of the new dispensation were that educational institutions had to create a new educational framework which was welcoming to all students irrespective of their background, as well as selection procedures which were equitable.

1.2. Focuses of this Study

This study focuses on the selection practices at the Wits Business School (WBS), which traditionally based its assessment of students' potential on a battery of psychometric tests. These were applied in addition to the use of indicators based on previous scholastic and academic performance as well as indicators based on previous working experience.

Enrolling into most South African postgraduate degrees requires one to obtain a certain average score. For example, 65% is the minimum average mark required for admission into most postgraduate degrees at the University of the Witwatersrand (University of the Witwatersrand Website, 2010). However for courses like the Master of Business Administration programme (MBA) there are additional requirements that students have to meet before they achieve entry into the programme.

At the WBS it is mandatory that prospective students are in possession of an undergraduate qualification and also have three to four years of work experience, preferably in a management position (WBS Information Brochure, 2007). The use of duration of employment as a critical

element in the selection criteria is a common element in South African MBA programmes despite the fact that there is no empirical evidence supporting the claim that work experience is related to academic performance in traditional MBA programmes (Koys, 2010).

Furthermore, applicants into the WBS MBA programme are required to undergo a selection test (WBS Information Brochure, 2007). It is this practice which forms the departure point for this study. One reason for this is that the use of "selection practices have come to be regarded with suspicion as it is recognised nowadays that advantaged applicants tend to have more of the requisite skills in place than their disadvantaged counterparts" (Tayler, 1996, p. 4). Despite this argument, the use of selection tests as a criterion for admission into MBA programmes is a prevalent experience worldwide (Koys, 2010).

In many instances, psychometric tests have carried great influence amongst a set of selection indicators at university level (Koys, 2010; Lavin, 1965). This study aims to establish whether empirical evidence supports this.

1.3. Aims and Rationale of this Study

In the South African legislative and constitutional context post 1994, empirical studies are necessary to establish that methods used for selecting students for courses are appropriate (Buntin, 1994; Foxcroft & Roodt, 2001). This means that in order to ensure that the tests are appropriate for everyone who is expected to take part in them, practitioners need to conduct consistent research into the psychometric properties (reliability, validity, fairness and bias) of the tools used for selection purposes (Code of Practice, 1998).

In the WBS context, the need for this type of empirical study is sparked by the fact that the majority of Black students who applied and wrote the selection tests were however not accepted into the MBA programme (Z.S. Koyana, personal communication, June 30, 2010). Instead the trend has been for more White students to be accepted into the programme. This disparity is indicated by the elevated number of White students in the current sample.

While the factors underpinning disparity in student enrolments are complex and in South Africa involve socio-economic as well as language and educational factors (Cockcroft & Laher, in press; Foxcroft, in press), the use of psychometric tests may also be contributing to this continuing disparity. There is therefore a need for this study, for the reason that the four NIPR tests used at the WBS were originally developed on a population of candidates with English as a first language (Beukes, 1969; Lewis, 1993; Lombard, 1975; B. L. Visser, 1977). Specific problems may arise when similar tests are administered to candidates who speak English as a second or third language. According to MacFarlane (2006), the two language groups process tasks differently due to the way they interact with the language.

It was due to the abovementioned concerns that the researcher decided to conduct the current study and attempt to answer the following research questions. These questions focus on the relationship between the NIPR developed selection test battery used at the Wits Business School and academic performance. Two academic performance criteria will be used in the investigation:

- Academic performance of WBS students in the ten core courses comprising its MBA curriculum.
- Academic success as indicated by whether or not WBS students ultimately obtained their MBA degrees.

Two academic performance criteria were necessary in this study for the reason that many WBS students perform well academically in their ten core courses in the MBA but do not ultimately obtain their degrees. The reason is that they do not complete their research reports which are also a requirement for graduation.

The research questions below refer to both these criteria. They also refer to four NIPR tests as predictors. These tests are drawn from the NIPR High and Standard Level Battery which are a suite of psychometric tests developed in the 1970's and applied widely during the apartheid era (e.g.,J.M. Taylor, 1978, 1980, 1981; S. Visser, 1981).

1.4. Research Questions

- Are there significant relationships between the four NIPR tests and academic performance in the WBS MBA curriculum?

- Do the first order correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?

- Do the multiple correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?

- Do the four NIPR tests predict academic success in the WBS MBA programme?

- Do the four NIPR tests predict academic performance equally for the different racial and language groups enrolled in the WBS MBA programme?

- Do the four NIPR tests predict academic success equally for the different racial and language groups enrolled in the WBS MBA programme?

1.5. Summary

This study examined the relationship between four NIPR tests developed during the apartheid era and two criteria. The first is academic performance in the WBS MBA curriculum, and the second is overall academic success in the WBS MBA programme.

This is an important area of research for the reason that current selection practices in South African higher education may still reflect assumptions which applied during the apartheid era. Another reason why this area of research is important is that there is heavy reliance on selection tests in many universities both in South Africa and internationally. These practices are based on assumptions that the instruments used are predictive of subsequent academic performance in the curriculum, as well as overall success in meeting all the requirements necessary for obtaining a university degree.

This study also examines race and home language as factors potentially influencing both academic performance in the curriculum as well as overall academic success in the WBS MBA degree. The reason for this is that four NIPR selection tests used as predictors in this research were developed during the apartheid era in South Africa, on a White only or English first language speaking group. It may be that the circumstances which applied in this era of legislated segregation are now different in a post-apartheid and democratic era, and that these different contextual factors influence whether or not continued use of these particular selection tests is justified.

2. Background and Context of the study

2.1. The Context of the Study

From an economic standpoint, the democratisation of South Africa has meant that the country is now able to re-enter international markets and compete with other developed and developing countries for the same economic resources (Bezuidenhout, 1997; Doody, 1997; Mulder, 2004). However re-entering global markets has meant that the country has exposed itself to the pressures and complexities of internationalisation (Bricker, 1990).

Changes in these global markets are likely to positively or negatively impact on South African's economy. Bricker (1990) suggested that South Africa was already experiencing increasing unemployment in the apartheid era. The lifting of currency exchange regulations and technological changes due to global economic pressures may accentuate this problem.

Bricker (1990) further emphasizes that these economic challenges have imposed pressures on business schools to create programmes that can develop managers with the necessary skills to take decisions enabling South Africa to compete at the international level. The demands of globalisation and industrialisation also call for different kinds of employees and managers (Carmichael, 2003; Strasheim, 1998; B. L. Visser, 1978).

One of the sectors that are likely to be called upon to meet the demands for a new type of manager is the education sector, with a concomitant elevated demand on higher education (Buntin, 1994; Fraser & Killen, 2003; Nzimande, 1995). In order to address the country's need for managers equipped with the necessary skills to assist with the economic challenges,

management programmes like the MBA degree have been created. In recent years, the number of MBA degrees as well as MBA students has been on the increase (Bezuidenhout, 1997 Carmichael, 2003; Tlakula, 2009). This reflects an increasing demand for management skills as well as an increasing pressure on students to take higher degrees in the management area.

2.2. What is a Master of Business Administration?

The MBA degree is an advanced qualification that people attain in order to enhance their opportunities in work environments or in society in general (Carmichael, 2003; Kagan, 2003). Obtaining the qualification is more closely associated with the self-esteem needs (or even self-actualisation needs in some instances) than with basic needs. "The more education one has, the more one earns" (Ratangee, 2006, p. 7). Wiggan (2007) referred to this as the "achievement ideology"- the general belief that doing well in school will improve one's life chances and one's social mobility. Earning power is also aligned with the esteem need.

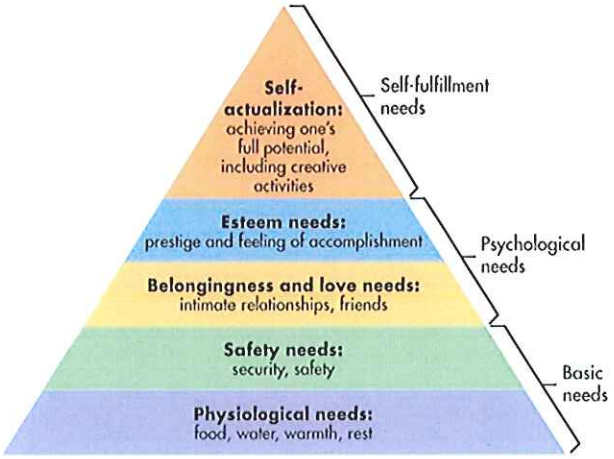


Figure 1: Maslow's Hierarchy of Needs

2.3. MBA Courses

According to Lewis (1993, p. 1) the objective of the MBA degree is "to prepare high potential individuals to lead and manage in complex and challenging environments of South African business". Roth (1989, in Bricker 1990) has said that the MBA degree gives students the skills to run a successful business, consequently to become good managers. The programme does not build the capacity to be responsible for an enterprise (entrepreneurship).

The Wits MBA curriculum includes principle courses, core courses and electives. Students are required to choose five elective courses from a selection of over thirty courses (WBS Information Brochure, 2007). These may be stream lined along the following areas: Finance, Information Management, Human Resources, International Business, Strategy, General Management and Marketing (WBS Information Brochure, 2007).

This curricular structure implies that the WBS MBA curriculum includes a variety of prescribed courses (See Appendix A), collectively aimed at helping students to understand all the important areas in their organisations (WBS Information Brochure, 2007). Accordingly, the qualification gives students generalist knowledge of organisations irrespective of the five electives they choose for specialisation.

"Embarking upon an MBA degree (with a background of at least two or three years managerial experience) enables students to rapidly develop a working knowledge of how specialist business domains (such as Business Transformation, Finance, Human Resources, Information Technology and Management, Marketing, Operations and Strategy) can work together for the good of the company" (Kagan, 2003, p. 7).

Although there is no empirical evidence to confirm that years of work experience play a pivotal role in the prediction of academic success, at the WBS this information is critical in the classroom setup where the depth of discussions is enriched by people with more corporate exposure (Cook & Owens, 1991, Kagan, 2003). These researchers argued that the less work experience a student has the less he or she is able to contribute towards classroom debates.

2.4. Costs of Studying for an MBA

Any form of education is a costly investment (Bricker, 1990). However the MBA degree is probably the biggest investment a person can make in him or herself (Weavind, 2011). That is why researchers like Carmichael (2003); Kagan (2003) and Yang and Lu (2001) stated that financial and non-financial expenses of studying for an MBA degree are high.

In 2003, in South Africa the cost of the best MBA degrees ranged between R20 000 and over R100 000 (Kagan, 2003). Current fees are R120 000 plus at some schools (WBS Information Brochure, 2007, University of Cape Town website, 2011). In the USA the fee was $87 500 for the top executive MBA degrees in 1999 (Yang & Lu, 2001). According to Weavind (2011), MBA degrees at the top US institutions like Harvard are likely to cost individuals and their respective companies around R1.2 million for a two year degree. This is very expensive, and what is more disturbing is the fact that often these fees exclude additional costs such as books, photocopying, accommodation and transport costs to and from campus (Carmichael, 2003; Kagan, 2003). Hence the total costs of studying for an MBA degree are likely to be significantly higher than what has been alluded to.

In addition to the abovementioned costs of studying for an MBA degree, non-financial costs such as sacrificing family time, reduced social contacts and elevated stress levels are also

worth considering as additional non-quantifiable costs of studying for an MBA (Carmichael, 2003; Doody, 1997). These variables have been shown to have detrimental effects on people's academic performance (Bezuidenhout, 1997; Carmichael, 2003; Kagan, 2003; Tlakula, 2009). Quite often MBA students do not have financial sponsors for their studies and most individuals in the age-range of the MBA class have families to support. Therefore, huge sacrifices are made in terms of time spent with the families and family finances are also affected.

The associated high costs of management education (more so the MBA degree) have contributed towards the need for firm selection criteria. The main aim of such criteria has been to exclude those students who are unlikely to meet the standard (the so called 'at-risk students') (Kagan, 2003; Killen, et al., 2003; Kotze & Griessel, 2008). These authors further argued that selection is justified on financial grounds for the reason that it assists in avoiding any secondary costs that come with failure to succeed in a programme. Killen, et al. (2003, p. 147) for example, state that "when universities admit students they need to be reasonably confident that those students will be capable of successfully completing the course in which they are permitted to enrol. To knowingly admit students who, for whatever reason have no chances of academic success is immoral".

As a consequence of a need for firm selection criteria, there has been growing interest in the measures used for the selection of students and a demand for the use of psychometric assessments (Foxcroft & Roodt, 2001; Nzimande, 1995). "When used properly tests can provide broader and more equitable access to education, training and employment with individuals, institutions and society benefitting" (De Beer, 2005, p. 6).

At this point many academic institutions use testing for selection purposes. The issue is not so much whether South African academic institutions are involved in selection of students, as

opposed to whether the measures meet the requirements as stipulated by South African legislation.

2.5. Implications for the Current Study

Due to concerns around high financial costs associated with studying and completing an MBA degree, it appears that selection practices are likely to be a permanent feature of the South African higher education landscape. One reason for this is that universities are determined to safeguard against accepting students who are not going to complete their qualifications and to only accept those students with high potential for completing their studies. While individual institutions have their own criteria for ensuring that they obtain high quality students, the South African education system is highly legislated. This is supported by increasing regulation of academic activities contributing to a context which prohibits academic institutions from refusing students based on minimum criteria (Buntin, 1994, Koch & Foxcroft, 2003).

In the apartheid era, the selection criteria adopted by many South African academic institutions involved but was not limited to psychometric tests. An example of these practices is found in the study by Potter (1991), which provides a case study of the development of a pre-university project in the Engineering Faculty at Wits which also linked with a project in the university's Management Faculty. The selection procedures adopted in these projects involved use of scholastic performance indicators, biographical inventories and psychometric tests. The four NIPR tests examined in this study were an integral part of the instruments used by the psychologists who developed the selection procedures.

While the current study focuses on selection practices in the WBS MBA programme, the results of the analysis have implications for the selection practices of other South African higher

education institutions. The reason is the majority of the NIPR psychometric tests were modelled on other tests developed in America and Europe, being normed on populations of primarily White candidates. The extension of this model cross-culturally can become a problem when the same tools are used (as in the South African context) on Black and White students alike without considering the background and uniqueness of the groups involved.

2.6. Summary

This chapter has focused on literature relevant to use of the four NIPR psychometric tests which formed the focus of this study. The literature review has stressed the aims of business school degrees, which provide the particular context for which selection procedures into MBA programmes are evolved. Critical to the use of psychometric tests in selection into institutions of higher education would be appropriate and justified use of the instruments.

The remaining chapters of this research report focus on whether the use of a particular suite of four NIPR tests in the WBS MBA programme can be justified empirically. Available data from student records at the Wits Business School will be examined to establish whether or not the four NIPR tests which formed part of the MBA selection process over a number of years were actually able to predict academic performance and also whether the scores yielded by the tests were able to distinguish between students who were likely to succeed academically in terms of actually getting the MBA degree.

Race and language have also been included in the analysis for the reason that many of the students identified in the selection process as unlikely to succeed were Black students. While it would have been ideal to have also followed up the success of those who were not selected for the degree, this was not possible within the scope of a Masters level research report.

Nevertheless, from the data available from the Wits Business School's archives it is believed that there is sufficient evidence of numbers of students to address the central concern as to whether the use of the four NIPR tests in the selection test battery was able to provide the type of information those involved in the selection process would have needed to make an empirically justified selection decision.

This is the central thesis examined in this research report. Another central thesis is whether those involved in the selection process had the data necessary to be able to adequately take account of racial group and language differences. This latter concern stems from the reality that during the apartheid era in South Africa, the four NIPR tests which form the focus of this study were developed and normed on groups of White students and the tests were subsequently applied for normative purposes to other racial groups. The argument would be that there may be a bias in the ways in which particular instruments are developed. It may also be that the types of selection practices developed in apartheid days are not the types of practices which are necessary in a post-democratic era in this country.

For these reasons the literature survey presented in this research report has three parts. In this chapter the focus has laid on the context of the WBS MBA programme for the reason that the analysis presented in the rest of this study will focus on the four NIPR tests used at the WBS in the selection of potential MBA students. In the next chapter a broader literature review on the use of cognitive and non-cognitive testing in higher education will be presented. Research on the psychometric properties of this particular suite of four NIPR tests will then be reviewed in chapter four.

The researcher will use this three part applied psychological literature as the conceptual basis for examining empirical evidence drawn from the Wits Business School (WBS) MBA archives concerning the use of the four tests in the particular context of the WBS selection process.

3. Literature Review

3.1. Introduction

Research into the predictive validity of pre-admission measures has received extensive attention over the years (Killen, et al., 2003; Koys, 2010). In the South African context, this popularity is associated with the increasing diversification of the student population. The consequences of these increases were competition for limited resources among the different groups (Buntin, 1994). Limitations on the capacity to accommodate all the students applying to the institutions of higher education has led institutions to seek valid and reliable tools through which they can determine students who are likely to complete their tertiary studies (Killen et al., 2003).

However as Lavin (1965) mentioned, it is not easy to conduct prediction studies and accurately compare students if they do not take similar courses during their studies. Students major in different subjects and some majors are known to be more difficult than others (Lavin, 1965). An illustrative example of this is between two Bachelor of Commerce (BCom) students from the same institution in the same year of study one majoring in Accounting and the other in Business Management. Although both students studied B Com, the level of difficulty encountered in studying Accounting might not be the same as that experienced by the student taking the Business Management course. Similarly within the current study, students choose electives based on their preferred stream of specialisation or the level of ease/difficulty associated with the course (Z.S. Koyana, personal communication, June 30, 2010). Therefore comparisons of student performance based on preferred choices are made difficult as discussed within the paragraph.

Also important in prediction studies are the criteria used by educators in assigning grades (Lavin, 1965). There is a degree of subjectivity inherent in the grading process especially for those disciplines that do not have a single straightforward one-word answer (for example, Social Sciences compared to the Natural Sciences and Mathematics). Quite often in the Natural and Physical sciences there is a straight answer to a question with minimal room for subjectivity on the part of the rater/educator (Lavin, 1965). Often in Social Sciences responses to questions are in essay or paragraph formats and the raters use their own judgement to decide whether the question was answered correctly and how much the response given is worth (in terms of allocation of marks). Thus, in such courses there is a lot of subjectivity from the rater's part.

Lavin's last points on the dilemmas associated with prediction studies was that very little has been done in controlling the variations as discussed above, therefore it is essential to be cautious when comparing students according to grades.

Further debates into prediction studies focus on the comparability of degrees from different countries. According to Ardilla's study "a Columbian professional degree can be viewed as equivalent to a Master's degree obtained in a US University" (2001, p. 412). Therefore comparability will be difficult to achieve based on the variations between degrees from these countries. A proposed solution is to always compare countries that have similar education systems. However this is not always feasible considering that students choose to study in different countries and some of these countries will be using a different education system to what one is used to.

The influences of qualifications from different countries is a concern at the WBS because the school tends to receive and accept applications from international students mainly from other African countries, but as well from European and America. Luckily in South Africa the process of

streamlining foreign qualifications is done by the South African Qualifications Authority (SAQA) (Buntin, 1994). However even though comparing degrees from different countries is addressed through SAQA, there is a possibility of experiencing the effects of diversified qualifications in the classroom situation.

The abovementioned points are very critical when determining admission criteria more so in postgraduate studies. At the WBS students come with a variety of qualifications, and if similar, with different majors, and qualifications from different countries. Therefore precautionary measures need to be taken when administering prediction studies with such diverse populations.

Considering that psychometric tests are used extensively in prediction studies it becomes necessary to address issues pertaining to the reliability, validity, fairness and bias associated with the psychometric measures.

3.2. The Importance of the Employment Equity Act in Prediction Studies

The use of Psychological tests in South Africa is guided by the Employment Equity Act (No. 55 of 1998), herein referred to as The Act. According to Section 8 of The Act, the use of psychological tests are prohibited unless it can be determined that tests in use are indeed measuring an aspect of behavior consistently (**reliably**), secondly, whether each test is measuring what it claims to measure (**validly**), and lastly, whether the test measures what it claims to measure in the same way across different groups and people (**fairly and without bias**) (Huysamen, 1994; Lopes, Roodt & Mauer, 2001; Murphy & Davidshofer, 2005; Urbina, 2004).

Historically more emphasis was placed on ensuring validity and reliability in measures (Paterson & Uys, 2005). Issues pertaining to fairness and bias started taking center stage with the increasing diversification of institutions and the democratization of South Africa. Before this, tests would be administered to people without taking their background factors into consideration.

The reliability of measures is a well-researched subject with most studies on properties of psychometric measures tapping into reliability as a first point of reference (Foxcroft & Roodt, 2001; Huysamen, 1983; Murphy & Davidshofer, 2005; Rosenthal & Rosnow, 1991; Urbina, 2004). "A test's reliability can be established internally by calculating the inter-item correlations or over time by examining the stability of tests scores" (Israel, 2006, p. 10).

Critical to the use of psychological testing is determining the validity of the measures for use in research. This means finding out the degree to which they measure what they are supposed to measure or claim to measure (Cronbach, 1960; Foxcroft & Roodt, 2001; Rosenthal & Rosnow, 1991; Urbina, 2004). Validity involves accumulating evidence in three categories: content related validity, criterion related validity and construct validity (Foxcroft & Roodt, 2001; Leedy & Ormrod, 2005; Rosenthal & Rosnow, 1991). Of importance to the current study is the predictive validity of a test, "the accuracy with which a measure can predict the future behavior or status of an individual" (Foxcroft & Roodt, 2001, p. 50). This is an area where psychometric tests gain popular use.

Predictive validity is a major form of statistical validity and is used whenever tests are used to make predictions about a specific outcome (Rust & Golombok, 1992). However a problem with predictive validity is that individuals who are not selected do not have a scorable criterion variable, so it is often argued that data from such studies are always incomplete. In order to

confidently confirm that a particular variable predicts or does not predict academic success it is vital that all people who have gone through the psychometric tests are accepted. Then a proper decision will be taken about the abilities of the test/s to predict.

3.3. The Cognitive and Non-Cognitive Variables as Predictors of Academic Success

Several authors have noted correctly so, that there are far more important variables in determining academic success than those used regularly in admission practices (Payne, 2008). They agree that academic success predictors consist of cognitive measures pertaining to mental ability or intelligence and non-cognitive measures especially personality traits, attitudes, levels of motivation and locus of control (Culverwell, 1989; Ridgell & Lounsbury, 2004). Ideally, predictive studies should incorporate a review of both cognitive and non-cognitive variables as each one of them is responsible for a certain proportion of the variance in academic success (Mulder, 2004; Randsell, 2001; Ratangee, 2006; Tayler, 1996).

3.3.1. Non Cognitive Variables

Non-cognitive variables such as personality, motivation, biographical variables, socio economic status, interests, self-concept, educational, cultural background, family and peer support, and financial stability are the factors that individuals interact with (Fraser & Killen, 2003; Mulder, 2004; Ratangee, 2006). Various studies have looked at the importance of these variables in the prediction of academic success (Kaufman, 2003; Mokone, 2008; Nurm & Arnola, 2001; Payne, 2006; Randsell, 2001). Though not a focus of the current study, it is important, more so for future studies to take cognisance of these variables.

The common findings among the abovementioned studies was that non-cognitive variables were better predictors of academic success for Black students, while cognitive variables predicted better for White students. Potter's 1991 study of Engineering students found that in North America, Britain and Australia, a positive predictive relationship was present between cognitive variables and academic performance. Similar findings were also achieved in the South African context among White Engineering students (Kaufman, 2003). However the literature indicated that cognitive variables were not consistently predictive for students from minority groups (Potter, 1991), and the prediction of academic performance and academic success among the minority students was better explained by the non-cognitive variables (Lavin, 1965; Potter, 1991). This point is critical in the current study as it also looks at the comparison of prediction results for different racial and language groups. It will be interesting to find out if the results mentioned here (tests predict better for White students than Black students) are also evident in the study however non-cognitive variables will not be dealt with.

3.3.2. Cognitive Variables

Cognitive variables are those variables that require one to rely on psychological processes to make sense of their environment (Eysenck, 1993). "These processes include attention, perception, learning, memory, language, concept formation, problem solving and thinking" (Eysenck, 1993, p. 1). The cognitive variables (like final year examination results, IQ scores, language proficiency, mathematics proficiency and aptitude tests) have dominated psychometric testing research for over 50 years (Koch & Foxcroft, 2003). They carry some importance in the prediction of academic success with more emphasis placed on the use of matriculation results and the standardised tests as predictive measures (Koch & Foxcroft, 2003).

3.3.2.1. Final Year Examination Results

"Worldwide, final year results from high school have generally been utilised for the purpose of admission into higher education" (Mulder, 2004, p. 19). Although other factors would become critical in terms of entry into institutions of higher learning (e.g. A Higher Grade pass in Mathematics and Physical Science for entry into an Engineering course at the University of the Witwatersrand) an exemption pass at matriculation still remains the essential entry point into University (Mokone, 2008). This was based on research findings that concluded that scholastic achievement was the best single predictor of academic success.

Researchers in Culverwell's (1989) and Shochet (1986) reports stated that school performance remains the single best predictor of academic performance however the vast amount of variance remains unexplained. Shochet's (1986) study found that correlations between matric rating and the average mark for the December exams for 52 Department of Education and Training (DET) students was 0.15 (p=0.30). He then argued that it is not possible to rely on results like these for selection purposes. Therefore other criteria for fair and accurate selection will need to be developed as it is clearly evident that that matric results are not the best predictor for Black students.

Despite the differences in education systems, South African institutions of higher learning relied heavily on matriculation results as a determinant of academic success (Buntin, 1994; Mulder, 2004; Nzimande, 1995). Shochet (1986, p. xvii) argued against this and said that "disadvantaged students have not had the educational opportunity to fulfil their potential at school and their school results cannon be regarded as indicative of their future academic success". His study found and concluded that matriculation results were invalid predictors of university success for disadvantaged students. Shochet (1986) made a very important point

about the use of tests in predicting academic success, which might bear consequences to the current study. He argued that only in exceptional cases have researchers been able to explain more than 25% (a quarter) of the variance in university grades.

J.M. Taylor (1978) conducted further studies on the subject and concluded that scholastic achievement was the best predictor of first year Engineering success at the university. He did however acknowledge the contribution made by psychological tests, even though the contribution was small.

3.3.2.2. Standardised Tests

Studies focusing on the Graduate Management Aptitude Test (GMAT) have established that the test is the best single predictor of grade point average (GPA) in first year graduate management programmes and numerous other types of programs (Elkin & Sneyd, 1989; Kagan, 2003; Paolillo, 1982). In Paolillo's (1982) study GMAT was the second best predictor with r=0.26 adding 0.04247 of the variance in academic performance. The primary predictor was undergraduate grade point and it explained 0.12311 of the variance. Therefore, the magnitude of the test though significant (0.001) it was small. In Fish and Wilson's (2007) study the total variance was less than 19% that too was very small. It was not one fourth of the variance in academic performance .Therefore a significant (81%) proportion of the variance in academic performance was explained by unknown variables. Thus heavy reliance on such regression findings would seem to be inadequate because a large amount of variance was left unexplained (Paolillo, 1982).

Yang and Lu (2010) also found that the undergraduate grade point average (GPA) and GMAT quantitative and verbal scores had a very significant impact on academic performance with the

r^2 of the regression at 0.26. Meaning more than one quarter of the variation in the MBA's academic performance could be explained by the regression model. Similar to Paolillo's study, undergraduate GPA was the most significant predictor of academic performance followed by GMAT. The exception in Yang and Lu's (2010) study was that they also included language as a predictor variable. However results indicated that language made little predictive contribution.

Wright and Palmer (1994) ran a different kind of prediction study. Theirs aimed to predict a dichotomous outcome (moderately low and very low performers). In order to achieve that they relied on a sample of 86 MBA students and found that GPA and GMAT were modestly associated with graduate performance but they did not discriminate between the moderately low and very low performers in the programme. r^2 was estimated to be 0.212. This point is very critical and relates very well to the current study especially on the section that looks at the ability of tests to predict whether one will successful complete or fail to complete their MBA studies. It is hoped that the four NIPR tests will discriminate between those who will be able to successfully complete their MBA studies and those who will not.

Irrespective of concerns raised against the use of matriculation results and standardised psychometric tests, psychologists continued with their reliance on these tools to make critical decisions about career or job fit, admission into studies and even diagnosis (Foxcroft & Roodt, 2001). A critical question to ask would be whether the tools used are reliable, valid, fair and unbiased. This is more important in the South African context, considering its history.

There are concerns about the ability of the GMAT test to predict academic performance for all students irrespective of race or language but this has not been adequately researched (Koys, 2010). Yang and Lu (2010) found that language was not a significant predictor. It is important in the South African context to find out the impact of language and race on prediction studies

considering that several management schools are using the GMAT as the admission criterion for their applicants.

3.4. The Role of Language in Predicting Academic Success

According to Cooking and Mestre (1998, as cited in Magadla, 1994), it has not been easy to understand the relationship between language skills and academic performance. Macfarlane (2006) argued that in order to achieve this it is important to understand that there are two types of language skills, namely, basic interpersonal communication skills and cognitive academic language proficiency. Quite often second-language speakers can come across as fluent in speaking English (interpersonal communication skills) but their performance on academic tasks (cognitive academic language) might not be as strong as those of first-language speakers.

Cognitive academic language is strongly related to reading abilities. Broom and Doctor (1994, p. 219) described reading as "a complex cognitive skill that requires the integration of many sub skills. Should any of these sub skills fail to develop adequately, an individual's reading ability may fall below expectation". Spoken language is cognitively undemanding but reading and written language requires far greater concentration and understanding (Macfarlane, 2006). "The majority of South Africans speak languages and dialects quite different from standard English as their mother tongue. It is therefore reasonable to assume that some of the generally observed discrepancies in test scores between White and Black South Africans are attributed to the variety of languages and dialects that are spoken" (Jensen, 1980 in Lopes et al., 2001, p. 62). Language serves a critical function in enhancing learners' academic success (Celliers, 2000 in Dockrat, 2005). It serves as a vehicle that enables learners to successfully complete their studies (MacFarlane, 2006). As a result, performance on assessment measures could be related to language difficulties rather than the abilities that the individual portrays. In practical

terms, the lack of language proficiency in the medium of instruction means that students will not be able to meaningfully talk, read and write nor to listen and comprehend what has been taught or substantiate a viewpoint. (Engelbrecht, 1999, as cited in Dockrat, 2005). Consequently students' inability to express themselves adequately in the medium of instruction (English) has a negative effect on their academic performance (Dockrat, 2005).

Unfortunately in the South African context students often learn in a language that they are not fluent in. The South African apartheid system "enforced mother tongue instruction for Black students with a systematic reduction of the role of English in Black education. This was also supplemented with an increasing emphasis on Afrikaans". (MacFarlane, 2006, p. 5). The advantaged students however had an opportunity to learn in their mother tongue throughout their schooling. Thus "the language segregation that resulted from the apartheid education system was also coupled with large scale social segregation so that Blacks were limited not only in their education opportunities but their freedom of movement and association was also heavily curtailed" (MacFarlane, 2006, p. 6). Similar to Potter et al. (1984), they further mentioned that this limitation of freedom of movement and association effectively denied Blacks access not only to education but restricted their opportunity to interact with proficient English speakers at almost every level.

The impact of language on academic performance is a driving force behind the current study and it would be interesting to find out if Dockrat's (2005) and MacFarlane's (2006) conclusions apply to the current situation. Both researchers found that tests of language proficiency on second language speakers have not shown any relationship to academic performance. It was at this juncture that J.M. Taylor (1978) argued for ability tests like the Mental Alertness because they seemed to have a better relationship with academic performance than tests of language proficiency.

3.5. The Role of Race in Predicting Academic Success

The testing movement in South Africa has long been associated with the Biesheuvelian Psychology which dominated from 1946 to 1973 (Nzimande, 1995). Its' major assumptions were inherent differences amongst the various racial groups in South Africa and it is further argued that this was shaped by the apartheid methodology. The apartheid dispensation used test results to influence people's opinions about the differences between the races particularly with regard to their intelligence. As a result White would be classified as superior to Blacks, Coloured and Indians due to their high scores in intelligence tests. This comparison of races according to their intellectual capabilities was customary (Biesheuvel & Liddicoat, 1958).

Eysenck (1981) and Jensen, 1985, in Tayler, 1996, agree that different races have different IQs as a result of their inherited genes. Both researchers provided evidence that the 15 point IQ gap between American Blacks and Whites have not been narrowed by compensatory educational programmes. It is unfortunate because such results are often based on tests that were developed on a White European population but similar results are anticipated with a Black population.

In trying to explain the educational experiences of African American students many researchers focus on deficient oriented explanations ignoring the students' strengths (Wright, 2009). Although these deficits are negative and display negativity towards the African American ways of doing things, Wright (2009) acknowledges that the formal education system does not favour Black students. The way it is designed favours White Middle Class Western students more. As explained by Ogbu and Simons (1998, as cited in Wiggan, 2007) many ethnic minority students (i.e., African Americans, Mexican Americans, and others) develop an identity that is counter to mainstream societal values including achieving in school.

With regards to comparing Black and White student's performances on psychometric tests, differences between means of White and Blacks were mostly two standard deviations apart and the test reliabilities were low for language based tests. Hence, support to the statement that said psychometric tests are appropriate for individuals from a White, Western, middle class background. Potter (1991) and Kaufman (2003) found that tests as cognitive predictors of academic success were more reliable for White students compared to Black Students. Few more studies were conducted by different researchers and the majority of them confirm the findings above.

"Different race groups in South Africa have lived very separate lives" (Holburn, 1992, p. 14). The use of a common test for all applicants relies on the assumption that the candidates have all been exposed to the same developmental opportunities. This is unlikely to be true in the South African context. Here different race groups lived in different areas, attended different schools under different education systems and spoke different languages.

3.6. Theoretical Perspectives into Intelligence Testing

Traditional/Genetics and Environmentalists have been the dominating theories in understanding the group differences with regards to intelligence testing (Jensen, 1973; Tayler, 1996). The genetics approach attempts to measure intelligence along dimensions which are purported to constitute the fundamental structure of intelligence (Tayler, 1996, p. 13). According to this theory "intelligence is a recognisable attribute and is responsible for individual differences. Intelligence is essentially genetically determined. It is stable over time and cannot be modified by external efforts to improve or increase it and it can be measured by IQ tests because of its occurrence throughout a number of different tasks" (Tayler, 1996, p. 13).

The genetic perspective asserts that there are immutable hereditary differences that separate Blacks from Whites (Wiggan, 2007). Advocates of this theory propose that differences in achievement and social class are a result of genetic variation in intelligence (p. 313). Madhere (1995, in Wiggan, 2007) disputed this assertion arguing that student achievement is a socially constructed outcome that cannot be viewed as an innate ability. Gardner (1983, 1993, in Wiggan, 2007) further argued that intelligence is multidimensional and cannot be measured accurately by unidimensional ability tests.

The theoretical basis that intelligence is inherited has dominated the field of psychology for quite some time. Galton was among the first theorists who found that some people are able to achieve intellectual status despite environmental hardships (Tayler, 1996). Implications of this to the current study are very important because there are some students whom, irrespective of their social circumstances were able to achieve academic success. It raises questions of whether genes played a role in their MBA performance or was it the combination of factors.

Spearman's genetically based single factor theory dominated the field of intelligence testing for a period of time. The theory was explained through the g-factor. Sattler (1982, in Tayler, 1996, p. 15) criticised the genetics approach arguing that "the g factor is a statistical construct which explains the co-variance between test scores and that a consequence of this correlational definition has been that so called intelligence has come to mean what the IQ tests measure". Furthermore, the genetics theory fails to consider the importance played by the environment on mental or cognitive development.

The environmentalist approach does not disregard the existence of genetic influences on cognitive abilities completely (Jensen, 1973; Tayler, 1996). According to this approach people learn from their environments (prenatal, social, cultural, educational and personality attributes)

(Tayler, 1996). A popular example is the one of two identical twins raised in different environments, what Plomin and Defries (1980, in Tayler, 1996) referred to as the twin and adoption IQ correlations. If genes accounted for the higher proportion of the variance in intelligence test scores, it would be expected that these twins would obtain similar scores in intelligence tests. Their study however concluded that "more than half of the variance in IQ scores is due to genetic differences among individuals" (Plomin and De Fries, 1980, in Tayler, 1996, p. 21). There was evidence of non-genetic influences to their intelligence test scores. (Tayler, 1996).

The environmentalists have not been able to provide concrete evidence to base their theory. As such, Gauvain (1993, cited in Tayler, 1996), argued that it is important to pay attention to the environmental influences and equally the genetical predisposition (Tayler, 1996). This approach has been referred to as the interactionist theory. "It holds that although there may be significant genetic differences at the time of conception, the organism's development involves such complex interactions with the environment that the genetic blueprint so to speak, becomes completely hidden or obscured beneath an impenetrable overlay of environmental influences" (Jensen, 1973, p. 49).

Basically, this theory argues that it is not easy to differentiate between the role played by genetics and the environment in people's intelligence. It is rather the interaction of all these factors. People's interaction with these aspects of life somehow determines the kind of person they are likely to become but as well these interactions assist in determining what they are able to learn (Jensen, 1973). The less the interactions the more the likelihood that learning will be prohibited. Thus Children's learning is heavily influenced by their parents and environmental factors they interact with (Jensen, 1973).

Tayler's 1996 study on the predictive validity of the Apil B test categorised the dynamic assessments as another perspective into intelligence testing. However dynamic assessments are not only about understanding group differences as suggested by genetic and environmental theorists. Dynamic assessments also focus on new ways of approaching assessments, more so when assessing people from different backgrounds.

When a measure is administered to a person who does not share the culture of the people who developed it and the country in which it was standardised, the assessment practitioner should always consider that the test scores obtained may not be a perfect reflection of the person's true level of functioning because of the role of cultural factors (Foxcroft & Roodt, 2001). Thus, there is a need to move away from tests that look at a person's current abilities and to look at the person's learning potential instead. Tayler (1996, p. iii) mentioned that "the concept of learning potential has come into its own as a result of dissatisfaction with the results of the genetics and the environmentalist approaches to intelligence testing". By looking at learning potential assessment practitioners are trying to move away from the influences and biases associated with cultural and social factors (Huysamen, 1983; Nzimande, 1995; Tayler, 1996).

3.6.1. Static Assessment

Static assessment focuses on the student's current level of performance on school related tasks (Gaydon, 1988). There is no indication of the student's potential to learn or improve. The theorists who use this approach defined intelligence in an innate and immutable form determined by hereditary traits and represented by a single IQ score (Frisby & Braden, 1992; Tayler 1996).

Most psychological assessments are classified as static assessments seeing that they look at the person's abilities in a still format. Feuerstein (1979, cited in Tayler, 1996) stated that "the conventional tests of intelligence measure intellectual achievements that exist at the time of testing" (p. vii). They use these measures to predict what the individual will be able to achieve or do in future. Hence the assertion that past behaviours are the best determinants of future behaviour.

T.R.Taylor (1994, in Tayler, 1996) argued that over time these tests are unable to provide sufficient explanations about the processes that cause changes in people's abilities and knowledge, as these assessments do not pay attention to the influences of environmental and social factors. If a person's assessment score says they are "below average", that will become a lifelong label that will not change, since static assessment does not recognise the potential to improve. By looking at the existing abilities, static assessments favours those students who have been exposed to better life conditions. Furthermore, the comparisons made through these tests are misleading as some ability tests have been found to have item and predictive bias (Jooste, 1995, in Tayler, 1996).

3.6.2. Dynamic Assessment

In his paper on Psychometric testing in South Africa, Nzimande concluded the article with some recommendations, one of which was that "testing will have to look at potential and not just actually existing skills in assessing people's capabilities" (Nzimande, 1995, p. 8). This sentiment was initially shared by Feuerstein (1979, as cited in Gewer, 1998), who said the object of focus for psychometric assessment should be the potential for being modified by learning.

The emergence of the dynamic assessment (DA) paradigm was prompted among other reasons by the recognition that the level of a person's current cognitive performance (often operationalized as problem solving) does not necessarily coincide with his or her learning potential (Foxcroft & Roodt, 2001; Tayler, 1996). Children from culturally or socio-economically disadvantaged back[g]rounds may not have been provided with the early learning experiences necessary to develop their potential" (Gaydon, 1988, p. 7). It is therefore important to look beyond pure screening but to also look at potential for future performance. According to Boeyens (1994. in Tayler, 1996, p. 10), learning potential is "the amount by which the performance of a testee will improve as a result of some instructional intervention".

Dynamic assessments focus on assessing what students are capable of achieving with instruction, in addition to what they currently know (Feuerstein, 1979, as cited in Tayler, 1996). The underlying philosophy of dynamic assessment is that the individual is an open system susceptible to influences that can produce structural changes in cognitive functioning (Jensen & Feurstein, 1987, in Gewer, 1998). This implies that "students from lower socio-economic groups are potentially as bright as their counterparts from higher socioeconomic groups, but because of their disadvantaged backgrounds, their current functioning does not match their potential" (Hoffenberg, 1988, p. 4). A child's academic status may not always reflect their potential. Children with poor academic status are often overlooked but other factors (like socio-political factors) could be influencing their performance (Hoffenberg, 1988).

Dynamic assessment is based on the notion that a zone of proximal development exists (Vygostky, 1987 in De Beer, 2005; Mulder, 2004; Tayler, 1996). Vygotsky further stated that the zone of proximal development is the difference between a child's capacity to solve problems independently and his/her capacity to solve them through mediation. It allows researchers to establish how much the student can learn from interaction with peers. A very crucial factor in the

WBS system is peer interaction, where students learn from each other through class discussions (Cook & Owen, 1991; Lewis, 1993). Lecturers are there as facilitators.

The on-going popularity of psychometric tests in South Africa can be attributed to their promise of providing means of fair assessment despite the vastly unequal and inadequate educational opportunities experienced by the majority of South Africans under apartheid (Sehlapelo & Terre Blanche, 1996).

Tayler (1996) alleged that a major advantage derived from dynamic testing is its relative lack of susceptibility to the effects of cultural bias. It can be easily used in those environments with high illiteracy or unequal living conditions (Tayler, 1996). Unlike static assessments, the comparison of individuals through dynamic assessments will not be dependent on their living standards and life exposure. Therefore, these methods have been shown to increase equity, fairness and accuracy when testing abilities and achievement in diverse student populations.

3.7. Summary

Racial differences in South Africa have been a focus of many studies more so in the behavioural sciences. The South African society has for years, been divided on racial lines. The institutions of higher education and industry were intimately involved in the apartheid process of racial segregation and racial division in terms of a context in which Black students found it difficult to access institutions of learning and were excluded from white collar roles in industry.

This literature review has been conducted from the standpoint that the history of the development of both cognitive and non-cognitive tests developed both outside and inside South Africa were developed on populations of White people and then also used on Black people. This

standpoint may be a biased one but if it is not, the alternative argument would be that the use of tests developed on White people for a highly diverse and multi-cultural population of South African students needs to be done carefully.

From this standpoint it would be important to acknowledge the effects of different cultures and environmental exposures on psychometric test performance. This is particularly the case in the legislative context of South Africa in which in order to rectify the effects of the previous regimes, the democratic government has regulated the use of psychometric tests in education and industry to allow for reliable, valid, fair and unbiased use of tests and test results.

The issue is that the use of particular tests needs to be empirically justified. It is also necessary to evaluate whether the empirical evidence justifies the fair and unbiased use of particular instruments in a context in which the majority of psychometric tests available commercially have been developed internationally and then adapted to the South African context. The argument would be that the adaptation process needs to be critically examined. Critical examination also needs to be conducted of the use of psychometric tests developed locally in South Africa, if these were developed under an apartheid dispensation and then applied on an on-going basis in a post-democratic era.

For this reason the literature review in the next chapter examines evidence on the reliability and validity of the four NIPR tests which form the focus of this study as background to analysis of the particular set of archival data on which this research is based.

4. The National Institute of Personnel Research (NIPR) Group Tests

4.1. Background

The popularity of group tests grew during and subsequent to World War 1 (Gregory, 2007; Huysamen, 1983; Nzimande, 1995). This was as a result of the need for mass recruitments of personnel into the army, and for the reason that one-on-one testing methods could not handle the increased demand.

Group assessments were seen as the solution because such tests are quick to administer to large groups with fewer test administrators (Huysamen, 1980, 1994; B. L. Visser, 1977). Quite often the administration process was also simplified and the process was uniform to all test candidates to allow even the less-trained staff to be able to administer the tests (Huysamen, 1994).

The major shortcoming of group tests is their impersonal nature which makes it difficult to collect data on test behaviour (Foxcroft & Roodt, 2001; Huysamen, 1983; Murphy & Davidshofer, 2005; Nzimande, 1985; Urbina, 2004). A multiple choice (forced choice) approach is the structure preferred for such tests. "Multiple choice tests force rote learning and penalise original, creative thinkers" (Murphy & Davidshofer, 2005, p. 89). The implication is that there is a danger that the multiple choice format limits a particular test to a simple exercise of rote memory.

There are also other potential limitations which relate to whether or not multiple choice tests may be reliable and valid. These stem from the reality that tests developed in one context and at

one time and shown to be valid in that context may not be as valid in another context or at another time. The implication is that test validation is a continuous process (Anastasi & Urbina, 1997). This study can be regarded as a step in that on-going process.

4.2. The NIPR High Level Battery (B/75)

The High Level Battery was developed as a group test by the National Institute for Personnel Research (NIPR) with the aim of being applied to applicants with twelve or more years of formal education (Beukes, 1969; Lombard, 1975; T.R.Taylor, 1987). It is a paper and pencil test designed to measure general intelligence, arithmetic ability and certain language abilities. (Beukes, 1969; Huysamen, 1983; Lewis, 1993; Lombard, 1975; Ngoako, year not specified; J.M. Taylor, 1978; B. L. Visser, 1977).

In addition to the selection of candidates, the High Level Battery can be utilised for vocational guidance and the classification of personnel (Beukes, 1969; Kendall, 1981; B. L. Visser, 1977, 1978). The battery has six subtests, namely, mental alertness test, arithmetic problems, English reading comprehension, vocabulary, lees en begripstoets and woordeskattoets (Beukes, 1969; Lombard, 1975; B. L. Visser, 1977). These authors further noted that it is not compulsory to administer the battery of tests simultaneously or in a specific sequence because the tests have separate norms, and can be administered and used independently of others.

It was for this reason that the WBS chose to use the Mental Alertness Test and the Reading Comprehension Test as part of their selection test battery. These instruments would be used in conjunction with other tests that did not form part of the High Level Battery. Norms for each subtest would be used to interpret the person's raw score in comparison to their reference group. The guidelines presented in the test manual would be followed, that each subject should

be compared to a norm group similar, in as many aspects as possible, to the people with whom he would have to compete in the occupation or course for which he is being selected (Beukes, 1969; Kendall, 1981).

The tests administered by WBS were as follows:

4.2.1. Mental Alertness Test

The High Level Battery Mental Alertness Test is classified as a test of general reasoning abilities with questions posed in verbal and non-verbal format (J.M. Taylor, 1978; Viedge & De Villiers, 1998). The test is written over 45 minutes, and in that time applicants are expected to answer 42 verbal items (Beukes, 1969; Lombard, 1975; J.M. Taylor, 1978; B. L. Visser, 1977). To succeed a candidate would require general intelligence, the ability to understand written English and the ability to apply the kind of problem-solving abilities gained at school (Beukes, 1969; Lombard, 1975). Knowledge acquired through the traditional Western educational system is a prerequisite for solution of the problems, which include reasoning in terms of verbal analogies, classification of abstract concepts, recognition of similarity between concepts and extrapolation to or correction of numerical or alphabetical series, (NIPR, 1975 in Lombard, 1975).

Items of the test also include numerical and letter series, verbal analogies and other problems requiring reasoning abilities (Beukes, 1969; Lombard, 1975; Ngoako, year not specified; J. M. Taylor, 1978; B. L. Visser, 1977). The test looks at the ability to apply previously gained knowledge to new situations (Lewis, 1993). It can therefore be classified as a static measure.

What is Mental Alertness?

Mental alertness is not a recognised component of intelligence. It is the lay term for reasoning potential (Dale, 2010). By reasoning, Sternberg (1982, p. 235, in Dale, 2010) referred to "a set of thought processes that concentrate on the structure rather than the content of organised memory during the process of information or retrieval and the subject's ability to answer a number of different types of items".

The most common kind of inductive reasoning problem is that of inducing structure, for example series completion, classifications, and analogies (Dale, 2010). These abilities are usually referred to as 'fluid' abilities as they resemble problem solving behaviours and form an integral part of intelligence from a lay point of view, for example reasoning logically and well.

4.2.2. Reading Comprehension Test

The High Level Battery Reading Comprehension Test is made up of twenty questions based on four different articles which took twenty minutes to answer (Beukes, 1969; Lombard, 1975; B. L. Visser, 1977). The test assesses the applicant's ability to understand written English.

Comprehension is a process by which the reader constructs meaning by interacting with the text (Anderson & Pearson, 1984, in Cooper, 1986). The reading comprehension process entails the reader encoding and decoding information and relating it to information that has been previously stored in their mind. Thereafter, readers construct their own meaning from within their own experiences (Cooper, 1986; Grellet, 1981).

Reading is essentially a language process (Broom & Doctor, 1994). Reading comprehension requires that the student develops the necessary skills to decode the language used in its written form and also requires that the student has developed appropriate oral registers of language against which to understand the content of the text. This explains the differences in reading comprehension performances of people with learning difficulties, people speaking English as their second or third languages, as well as people from different cultural groups on reading comprehension tests. The process of constructing meaning in relation to what has been read is thus influenced by both cultural and environmental circumstances (Cooper, 1986).

In the process of constructing and attributing their own meaning to the text, readers use what Cooper has called schemata - structures that represent the generic concepts stored in an individual's memory. Schemata theory explains how these structures are formed and relate to one another as an individual develops knowledge (Cooper, 1986) and also argues that a reader develops schemata through experiences. If the reader has limited (or no) experience with a given topic s/he will have insufficient (or no) schemata to recall and comprehension will be difficult or impossible (Cooper, 1986).

4.2.2.1. Registers of Language Used in Wits Business School Courses

The courses at the WBS require students to read texts and analyse business case studies written in English (WBS Information Brochure, 2007). These readings come with large amounts of material which need to be read and understood, so both reading speed and reading comprehension are important. Reading difficulties become most apparent when the reader is unable to grasp the meaning from a text passage (Broom & Doctor, 1994; Cooper, 1986).

He has also argued that attitudes towards reading affect the way one comprehends. Previously held attitudes about a subject are likely to influence how one comprehends a text. An example provided in a text for example, might reinforce or conflict with politically held views (e.g., a text describing a politician who has made a better life for himself). Someone with their own views about the topic might miss out nuances and other details of a text and relate everything to their own previously-held conceptions.

The literature on reading comprehension also suggests that reading comprehension may be affected by other factors such as the difficulty of the text, the vocabulary used in the text and the reader's familiarity with the subject matter (MacFarlane, 2006). What this implies is that performance on a reading comprehension test such as the NIPR High Level Battery Reading Comprehension Test used as part of the WBS selection process would be influenced by a number of language, experiential, attitudinal as well as cultural factors.

4.2.3. Arithmetic Reasoning Test (A/131)

Conventional numerical tests tend to measure the ability to perform routine calculations (Halstead, 1985). "The Arithmetic Reasoning test measures understanding of the fundamental rules of arithmetic and the application of these rules in logical strategies in order to solve problems of a purely numerical nature" (Halstead, 1985, p. 1). Items are presented in a novel format to prevent the use of rote arithmetic procedures. Also worth noting is that the items are cumbersome to solve algebraically, so that testees with algebraic experience have no advantage over those with little algebraic experience.

The high level battery arithmetic reasoning test has twenty verbally stated arithmetical problems that have to be completed in 40 minutes (Beukes, 1969; Lombard, 1975). Testees are expected

to abstract the problem from its verbal context and perform certain arithmetical operations to obtain a solution (Beukes, 1969; Halstead, 1985). "Knowledge required of the testee to solve the problems goes beyond the arithmetical skills and reasoning needed to choose the appropriate strategy, familiarity with a variety of concepts, measuring scales and units, objects and practices particular to a Western society is assumed," (NIPR, 1975, p. 6 in Lombard, 1975).

The decision was taken at the WBS not to use the High Level Battery arithmetic test but instead to use a standard level test that has 30 items and students had to attempt it in 30 minutes.

Questions in the arithmetic reasoning test consist of numerical equations which involve the determination of the value of the unknown digits (Halstead, 1985). The understanding of the fundamental rules of arithmetic and application of these rules in logical strategies in order to solve problems of a purely numerical nature are measured by this test. The test is useful in predicting success in activities involving the manipulation of numbers or symbols representing numbers (Halstead, 1985).

Several logical and numerical steps must be taken in order to solve each equation. The test presents the subject with numerical material in a format which he is unlikely to have encountered previously (J. M. Taylor & Halstead 1985 in Halstead, 1985). Hence, it is said the test measures understanding of the fundamental rules and operations of arithmetic and the use of these in logical strategies.

4.2.4. Conceptual Reasoning Test (1/0/1990)

This test involves logically deriving general rules from specific examples (Sternberg & Gardner, 1983, in Boeyens, 1990). That is reasoning from part to whole or from particular to general (Boeyens, 1990). It is for this reason that the test is considered a test of inductive reasoning.

It takes 35 minutes to answer 35 questions. The Conceptual Reasoning Test is a non-verbal measure (or performance test) of inductive reasoning ability in which candidates are required to identify the nature of the problem solving strategies to solve them (Boeyens, 1990; Viedge & De Villiers, 1998). This does not mean that the items cannot be solved using verbal strategies. "It merely means that testees with low verbal skills are not necessarily penalised for having low vocabularies in the test language" (Boeyens, 1990, p. 2). In contrast, he argued that the verbal medium of presentation in the Conceptual Reasoning test is problematic in South African where testees have a diversity of first languages.

Four distinct types of problems are clearly defined as series problems, two-way classification problems, transformation problems and operations problems. The emphasis in the Conceptual Reasoning Test is on the identification of concepts from a known (given) repertoire and the application of these to a wide variety of contexts rather than on the attainment of concepts from prior knowledge and experience (Boeyens, 1990; Holburn, 1991a). It should be emphasised however, that disadvantaged people are likely to score lower on the Conceptual Reasoning Test than advantaged people (Boeyens, 1990). The test constructors have tried to reduce this effect but have not eliminated it (Boeyens, 1990). It is therefore questionable why some institutions are continuing to use this test.

Many nonverbal ability tests contain a diversity of different types of items. The testee has to use several different strategies for solving the items in the test. Good education and a privileged socio-economic environment enhance the acquisition of a diversity of problem solving strategies (Boeyens, 1990; Holburn 1991a). Conversely, poor education and intellectually impoverished surroundings hinder the acquisition of effective problem solving strategies (MacFarlane, 2006). In other words socioeconomic and educational background will have an influence on the number and quality of problem solving strategies a testee brings into the testing situation. The number of problem solving strategies an individual has previously accumulated is the single most important determinant of test performance and knowing how to tackle a problem greatly reduces the problem's difficulty (Boeyens, 1990). It can be noted that an individual from a disadvantaged background may perform poorly on a nonverbal ability test merely as a result of inadequate previous learning opportunities.

Literature argues that nonverbal tests are ideal in situations where people from different cultural backgrounds are tested simultaneously (Jensen, 1973). It is believed that such tests address cultural bias because "they limit the bias introduced by requiring that candidates respond to test items in a second or third language" (T. R. Taylor, 1997 in Lopes et al., 2001, p. 62). There is no communication involved in processing these tests. However as argued in the chapter, some elements of bias are still observed with nonverbal tests as these tend to be heavily reliant on exposure and previous experiences.

4.3. South African Research on the Above NIPR Tests

4.3.1. Reliability of the Four NIPR Tests

There are two approaches which can be used for estimating the reliability of a test which do not involve having to administer the test twice. The first approach correlates results from one half of the test with the other half of the test, and this is referred to as split half (Gregory, 2007). The second approach examines internal consistency of the individual test items determining whether the items tend to measure the same factor (Gregory, 2007).

Values under 0.6 are regarded as indicating poor reliability and the tests are probably not measuring the same trait accurately or consistently (Galpin, 2001). Values above 0.7 are generally considered acceptable (Anastasi & Urbina, 1997). Internal consistency reliability estimates as indicated by Cronbach Coefficient Alphas, are presented in Table 1 below for each of the four NIPR tests, namely, Mental Alertness, Reading Comprehension, Arithmetic Reasoning and Conceptual Reasoning, as a measure of their internal consistency (Murphy & Davidshofer, 2005). These indicate that the four NIPR tests in this study display strong levels of consistency all at (0.84). The findings for Conceptual Reasoning are on par with Boeyens studies, where reliability coefficients in his studies were between 0.82 and 0.86, the only exception was 0.66.

Cronbach Coefficient Alpha		
	Raw Variables	**Standardised Variables**
Mental Alertness	0.833356	0.846456
Reading Comprehension Test	0.839401	0.852484
Arithmetic Reasoning Test	0.83985	0.852918
Conceptual Reasoning Test	0.836558	0.850109

Table 1: Reliability Results for the Four NIPR Tests

In addition to having adequate reliability, evidence concerning the validity of the four NIPR tests is presented in the sections following.

4.3.2. Literature on the Use of the Four NIPR Tests in Combination

A literature search did not yield many validity studies conducted on the four NIPR tests which are the focus of this research. This may have been due to the fact that the tests were originally developed in the 1970's, before computerization became the norm for organising data bases. It may also have been due to the fact that the NIPR was an autonomous body in the 1970's which then subsequently became part of the Human Sciences Research Council. It may also have been due to the fact that much of the NIPR's research was published as occasional papers as part of its own occasional paper series, which was organized on microfile as opposed to being linked into other computerized data bases.

Be that as it may, repeated attempts to access NIPR reports through libraries and through the HSRC library did not yield a long list of relevant research reports. Scrutiny of the tests manuals indicated that each test had its own set of norms but that these norms were mainly created from

White samples. The manuals also indicated that the tests were designed for group use but could also be administered individually. However, the researcher was unable to find a study that utilized all four tests simultaneously.

A study that almost achieved this was by Boeyens (1990), who undertook a correlational analysis on Mental Alertness, Reading Comprehension and Conceptual Reasoning. He concluded that a strong relationship (r=0.64) existed between Mental Alertness and Conceptual Reasoning Test. A moderate relationship (r=0.3) was evident between the High Level Reading Comprehension Test and Conceptual Reasoning Test. All these relationships were significant at 0.0001. He further confirmed that the relationship between Mental Alertness and Reading Comprehension yielded a 0.43 correlation coefficient.

However, most studies that focus on the use of any of the four NIPR tests in the South African context put more emphasis on the Mental Alertness Test. A study by J. M. Taylor (1978) for example, focused on first year Engineering Technician students in order to predict academic success amongst the group. The rationale for the study stem from an alarming failure and dropout rate of engineering students, therefore there was a need for a study that identifies factors that contribute towards this high failure rate.

J. M. Taylor's (1978) study used the Blox Test, Gottschaldt Figure Test, General Science, Mental Alertness and Deductive Reasoning to determine academic performance. His sample consisted of only White students who spoke English or Afrikaans. The findings from the study revealed that there were no major differences between the scores of the groups. Mental Alertness, though not a first best predictor it came out as the second strongest predictor of academic performance. He was quick to acknowledge that these findings should not be

generalized to all people as his focus was mainly on a White English and Afrikaans speaking sample.

4.3.3. Literature on the Use of the High Level Battery Mental Alertness Test

J. M. Taylor (1978, 1980, 1981) and S. Visser (1981) both used the mental alertness test as a control for intelligence in their studies of the relationship between three dimensional spatial perceptions and the academic performance of engineering students. Potter (1991) in his work with Jamotte of Anglo American used the mental alertness test in a similar way, and in addition created sets of norms for large samples of African matriculants. However these norms were used internally within the corporation and a number of affiliate companies to enable the award of scholarships, and were not published.

Lewis (1993) went on to administer the Mental Alertness on a Black student. She argued that the NIPR tests have a high reliability factor but are still prone to errors of measurement. She assessed two candidates using the Mental Alertness Test. These candidates had similar prior educational backgrounds, the only distinguishing factor was their race. The variation in the assessment results led her to conclude that "the differences in scores points to the possible problem of a language difficulty factor and cultural bias within the test administration procedure" (Lewis, 1993, p. 5).

Another study that focused on the Mental Alertness test was recorded in Saunders' 1993 Masters thesis. Her findings revealed that the test had the highest correlation with language proficiency and test performance. The study also revealed that the predictive validity of the Mental Alertness test for candidates considered for a bridging course at the Faculty of Economics and Management was not established and the mean scores in all the tests were

very low, i.e. 30% and below, with the Conceptual reasoning test being the only test which had a mean above 30% (37.9%). It would be interesting to see the effects of Mental Alertness of relationships with language. It is the aim of the study to find out if there are group differences on how the tests predict academic performance for different language groups.

The above studies have been cited to indicate that there may be further literature on the use of the four NIPR tests which the researcher was not able to access. The same limitation may have applied to the use of the four NIPR tests in the Wits Business School for selection purposes.

Studies on the reliability of the Mental Alertness Test were conducted at the WBS in 1982 and 1983, and revealed a Kuder Richardson (KR) 21 of 0.67 and 0.77, respectively (Cook, year not specified). The reliabilities for the English Reading Comprehension were 0.62 and 0.607 (Cook, year not specified). These indices are cited as they yielded relatively low internal consistencies. There was no information available on the research methodology used, the sample description or the analysis procedure utilised. In addition, the data sets are at this stage old, and as has been argued in Chapter Two, the contextual conditions applying in post-apartheid South Africa are different to those applying prior to the advent of democracy in 1994.

The current study thus aims to analyse more current data on the use of the four NIPR tests when used in the WBS context. Analysis of the reliabilities of the instruments has not been possible as the researcher has been working with archival data from student records, and the original test protocols are no longer available. Nevertheless, the assumption has been that correlational analysis of what are large data sets could yield findings of value.

The standpoint of the researcher is that South Africa has gone through wide-spread socio-political and economic changes since the 1990's. There is therefore a need for updated

research on the NIPR High Level Battery as well as other NIPR tests developed in the apartheid era, which may be useful for those institutions that are still utilising the NIPR tests for purposes of vocational guidance or as a basis for other types of decision making.

4.3.4. Literature on the Use of the Arithmetic Reasoning Test

The researcher was not able to find many studies which had been conducted using the arithmetic reasoning test. A few that were conducted were recorded in Holburn's (1991a) study and these are paraphrased below.

Holburn mentions a 1984 study of 107 White technikon and technical college students between the ages of 20 and 30, in which the test was found to have a 0.84 reliability coefficient. The mean was 17.02 (*SD*=6.8). A 1985 study of White male and female Bachelor of Science applicants revealed a reliability of 0.836 (which is equal to the one on technikon students). The researcher only found one study conducted in 1983 on Black Bachelor of Commerce bursary applicants. The reliability coefficient from this group was 0.84. But compared to other studies of University applicants in the same year, this score is lower than the reliabilities of 0.9 that most White applicants had (Holburn, 1991a).

Boeyens' 1987 study on the waitlisted BSc students at the University of the Witwatersrand found a correlation coefficient of 0.69 between Conceptual Reasoning Test and the Standard Level Arithmetic Reasoning Test (Boeyen, 1990). This confirmed a significantly strong relationship between the tests.

4.3.5. Literature on the Use of the Conceptual Reasoning Test

The researcher could also not find many studies that used the Conceptual Reasoning Test for prediction purposes. The studies that the researcher could find focused mainly on the test's reliability. Boeyens (1990) was one of the researchers who explored the reliability of the Conceptual Reasoning test against other measures, including the Mental Alertness and Reading Comprehension. He conducted these studies with High School, technikon and University students from the Black, Indian, White and so called Coloured races. In all instances highly correlated relationships between Mental alertness and Conceptual reasoning were observed. He reported correlation coefficients of between 0.64 and 0.68.

Relationships between Conceptual Reasoning and the High Level Reading Comprehension on the other hand yielded correlation coefficients of 0.3 (p=0.01) for the St Stithian study and 0.24 (p=0.05) for the 1989 study of 136 Promat pupils. These values, though significant were not strong.

It will be evident from the above that much of the literature on the use of the four NIPR tests which form the focus of this study is dated, as well as a little patchy. For this reason, the argument would be that the current study is useful, and that its findings could serve as a foundation for decision making concerning the continued use of the four NIPR tests for selection purposes at the Wits Business School. It may also form the basis for additional studies that look beyond the psychometric properties of the NIPR tests used at the WBS, to other forms of selection procedures.

4.4. Summary

In summary, since the first democratic elections in 1994, South Africa has experienced rapid changes in all spheres of society (De Beer, 2005). "These changes have also influenced psychological measurement and there are new needs identified in the changing and developing multicultural society" (De Beer, 2005, p. 1). There is a need for a study that looks at the performance and comparison of different racial and language groups with regards to psychological tests, for the reason that current students entering South African universities have a diversity of abilities and backgrounds (Killen et al., 2003, p. 147). Accordingly, there is a need for tools that acknowledge these differences.

5. Research Methodology

5.1. Introduction

This study focuses on four NIPR tests namely, Mental Alertness, Reading Comprehension, Arithmetic Reasoning and Conceptual Reasoning using available data from student records in the data base of the Wits Business School (WBS). The data were drawn from the original files stored in the WBS archives.

It was found in the process of compiling the data set that the High Level Battery Reading Comprehension test had been adapted by WBS staff prior to use. This study first addresses the issue of the relationship between the tests in the battery. As indicated in the previous chapter, the issue of the reliability of the Reading Comprehension test could not be addressed as the original test protocols were not available in the WBS' data base. This was a limitation of the data used in this study, which was not possible to address.

5.2. Research Questions and Hypotheses

Questions	Testable Null Hypothesis	Statistics
• Are there significant relationships between the four NIPR tests and academic performance in the WBS MBA curriculum? • Do the first order correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?	• There are significant relationships between the four NIPR tests, ten core courses and academic performance as defined by the average of the ten core course. • The first order correlations indicate that the NIPR tests predict academic performance in the WBS MBA curriculum.	• **Multiple Correlation**
• Do the multiple correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?	All four NIPR tests equally predict academic success in the MBA programme.	• **Forward Stepwise Multiple Regression**
• Do the four NIPR tests predict academic success in the WBS MBA programme?		• **Forward Stepwise Logistic Regression**
• Do the four NIPR tests predict academic performance equally for the different racial and language groups enrolled in the WBS MBA programme?		• **Forward Stepwise Multiple Regression** • **T-test** • **Anova**
• Do the four NIPR tests predict academic success equally for the different racial and language groups enrolled in the WBS MBA programme?	The four NIPR tests predict equally for the different racial and language groups in the MBA Programme.	• **Forward Stepwise Logistic Regression**

Table 2: Research Questions, Hypotheses and Associated Statistics

5.3. Assumptions

Being based on available data drawn from the WBS archives, it was assumed that the prospective students who wrote the NIPR tests were functioning adequately on the day of testing. This supports the assumption that the students were all in the same frame of mind when they took the test.

It was also assumed that students who were accepted into the MBA programme completed all their assignments and contributed equally to the syndicate assignments. This is more in line with the determination of academic success. Herein we assumed that students were actively involved in achieving their academic success and did not have other people doing assignments for them. This potential limitation in the data is mentioned, given that most participants in the WBS MBA programme were managers in their respective companies, there is a possibility that junior staff members could have been tasked with completing the assignments for them (Kagan, 2003). There is also a possibility that certain of the criterion academic performance data were contaminated by the WBS practice of having students conduct their work in syndicate groups, which would imply that marks assigned to an individual were influenced by (or even based on) the work of other individuals working in a particular syndicate group.

The study also assumed that none of the people writing the NIPR tests had a language barrier or difficulty when writing the tests. This assumption influenced the procedures applied in conducting outlier analysis in the data. What was found was that some English Reading Comprehension scores were low (more so for previously disadvantaged students). Nevertheless, the researcher has assumed that everyone in the sample of archival data available to her had a basic understanding of the English language and they could read the instructions and understand them.

Against these assumptions, the researcher has developed the following research design for this study.

5.4. Research Design

The most objective manner in which to evaluate a psychological test would be through research into the reliability and validity statistics of the test. It is argued that the type of research conducted in the current study is essentially evaluative (Potter, 2004), in the sense that evaluation is the systematic acquisition and assessment of information to provide useful feedback about some object (Trochim, 2006). The more specific and applicable for this study is formative evaluation, as it focuses on the process of programme implementation (Potter, 2004). It usually incorporates a process of programme monitoring to establish whether the intervention is being implemented as planned. In formative evaluation, the evaluator attempts to identify aspects of the programme which are working well, aspects of the programme which are problematic, and aspects of the programme requiring modification or improvement (Potter, 2004).

Similar assumptions underpin Potter's argument (in press) that in a developing world context, there can be value in considering assessment and evaluation literature as linked. One reason for this is that the evaluation literature offers a variety of contextually based approaches and models which can be of value in conducting the types of psychometrically based investigations which are necessary to extend, adapt or critique models developed in North America and Europe in relation to their wider use cross-culturally. Another is more mundane but perhaps equally necessary, that in using psychometric tests (or any other measurement) instrument in multi-cultural contexts, the use of these instruments requires justification. The process of justification is not only contextually but also evidence-based, implying an evaluative process.

A third reason is that there are intentions to use the results of this evaluative study to aid in decision making (Potter, 1998, 2002), concerning whether or not to continue to use the four NIPR tests for selection purposes for the WBS MBA degree. This study will thus provide recommendations at the end based on answers to the three research questions, with the aim of providing feedback which is perceived as "useful" in aiding decision-making (Trochim, 2006).

"The key assumptions of evaluation are firstly that programmes are initiated with goals in mind, they are therefore there to achieve those goals. Secondly, evaluations serve to collect information on the way in which programmes function and the effectiveness of these programmes in meeting the proclaimed goals. Finally and most important to the current study is the assumption on utilisation. Evaluation information will be used by decision makers in making discreet decisions which are likely to have an influence on the improvement of the programme" (Madaus, Scriven, and Stufflebeam, 1983, p. 178). The decision makers are the defined as the key stakeholders where evaluation is concerned. The term 'stakeholder' was proposed by Robert Stake in 1975, to refer to persons with different vested interests in a programme (Potter, 2004). Thus, the WBS is a key stakeholder in the current study.

5.5. An Evaluative Focus

The last point relates very well to the purposes of the current study. WBS management has shown an interest in the results of the current study as they believe it will help them decide on their admission criteria going forward (Z.S. Koyana, personal communication, June 30, 2010). It would thus be argued that management-oriented evaluation is the appropriate definition for the kind of evaluation conducted in this research. The study's findings will be written in such a way that they can be used to serve decision makers. The study will also be conducted in a manner that enables the information from the study to be utilised as an essential part of logical,

evidence-based making (Potter, 2004). The purpose is therefore to provide a knowledge and value base for making and defending decisions.

Theorists like Stufflebeam, Provus and Patton have played a critical role in the field of management oriented and evidence-based evaluation (Potter, in press). In this study, the researcher has employed the theory of Alkin (1969), who made decision making for programme managers the central focus of his model. Alkin defined evaluation as a process of ascertaining the decision areas of concern, selecting appropriate information and collecting and analysing information in order to report summary data useful for decision makers in selecting among alternatives (Potter, 1991). His CSE Model was like Stufflebeam's CIPP model (Stufflebeam, 1972) directed to enable decision-makers to use information to take decisions concerning educational programmes (Potter, 1991). This research focus is apposite to the evaluative aims of this study (Stecher & Davis, 1987).

5.5.1. A Non-Experimental Research Design Focused on Available (Archival) Data

The research design used in this study is non-experimental in that it is purely descriptive relying on archival data. The researcher collected all the necessary data from WBS archives in order to determine the association between scores yielded by the four NIPR tests and academic performance indices. The NIPR test data were then examined against an overall degree achievement versus non-degree achievement criterion. This was done with the researcher having no direct interaction with the process of data collection itself, and no direct interaction with any of the students on whose records this research has been based.

This research design is thus non-interventionary and non-experimental and based in a tradition of using unobtrusive measures (Webb, Campbell, Schwartz, & Sechrest, 1966). In terms of this

tradition, the evidence for this study has been collected from student records and the analysis of these data then conducted post hoc. In this non-experimental research design the independent variables were not manipulated, there was no control over extraneous variables and there was no random assignment of subjects into different conditions (Dockrat, 2005; Leedy & Ormond, 2005; Rosenthal & Rosnow, 1991). The researcher was essentially an observer (Webb et al., 1966), who aimed to quantify the relationship between academic performance and academic success as criterion or outcome variables and the tests used by the WBS as independent or predictor variables, to determine what level of correlation existed as well determine if the levels of relationship between the variables indicated a firm relationship between the independent variables, academic performance and overall academic success.

Unfortunately lack of control over extraneous variables means that the findings from this study may have been affected by a number of external factors outside of the researcher's control (Potter, 1998). There are therefore a number of potential threats to the internal validity of this study. This point will be discussed further later in this chapter as well as in describing the limitations applying to the analysis.

5.6. Procedure for Data Collection

Data stored in the WBS archives were taken as they were, tabulated and then analysed. Thus, an unobtrusive data collection approach was used. In obtaining data for this study the researcher had to first obtain permission from the university authorities. There were some practical difficulties in that the data had been saved by different departments within the university (Faculty of Management and University records). The admission test scores (the predictor variables) were stored by the Admissions Testing Office whilst academic course scores (the criterion variables) were kept as part of the University records.

Nevertheless, these different data sets were intact and once the necessary permissions had been obtained, the researcher was able to capture the data from these two different sources into one excel spread sheet. This was created in a way that enabled the matching of each individual's test data against their course results. As certain of the observations involved missing data, strategies had to be developed for how to deal with this problem, which are described in the sections following.

5.7. Sampling Methodology

The main aim of the study was to determine the relationship between the scores obtained by WBS MBA students on the four NIPR tests, academic performance and overall academic success. The reason for using two criterion variables in what was essentially a predictive study was that there was an additional influence affecting the criterion of pass or fail for this particular sample.

The MBA is only awarded to those students who have passed their coursework and also handed in their research reports. As it is an exacting programme, many students drop out. Many students continue to complete their coursework successfully but then do not complete the research component of the degree. There are many reasons for this, and to explore and document these lies beyond the scope of this study.

Suffice to say that from conversations with those who teach in the MBA programme as well as students who have been selected for and then taken MBA degree courses, there would appear to be a large number of reasons why individual students register and then do not complete their MBA degree. The cumulative effect of these third variables leads to a greatly reduced number of students who complete MBA studies. For reasons of the attenuation of the sample of those who

complete the degree (as opposed to those who started out with good intentions but did not stay the course), this study focuses only on those students who satisfactorily finished their coursework.

This is a limitation in sampling which at one level produced missing data relating to those students who registered and only partially completed the coursework component. The large cohort of students with only partial coursework marks would appear to have not continued with their studies due to factors applying to how this particular degree is structured, which affects many students.

The sample of student records with which this research deals is an attenuated sample of 741 students on who there were not missing data. It includes students who have satisfactorily completed their coursework only, as well as those who have satisfactorily completed their coursework component and then subsequently handed in their research reports. The sample of 741 students then splits into a further two components, 429 students who satisfactorily completed both academic and research components and then obtained their MBA degree, and 312 students who satisfactorily completed the academic component by completing all their ten core courses, but who then did not hand in their research reports.

Satisfactory completion of the whole degree versus satisfactory completion of only the academic requirement is thus treated as a variable in the research design. Eliminating those who had not submitted their research reports, would not only have had a negative impact on the sample size, but might also have limited the findings possible from this study. For this reason, the decision was taken to include the results of the group of 312 students who had achieved success in their coursework but had not necessarily been awarded their MBA degrees. This is mentioned as a

sampling factor as purposive sampling and sample attenuation introduce further limitations in the data and in this study as a whole.

In this study the data set was purposively structured to include the results of MBA students who had both graduated as well as results of MBA students who had achieved academically but had not graduated. It did not include the results of students who dropped out along the way and those who did not complete the academic requirement (as opposed to the research requirement of the degree.

This purposive sampling strategy was adopted based on what the study aimed to achieve (Dockrat, 2005; Huysamen, 1983; Leedy & Ormrod, 2005). This decision introduced limitations affecting both internal and external validity. Essentially, generalisability is limited because of the way in which data organisation and data cleaning were conducted in this study. The sample of data analysed may also not be reliable in the sense of representative of all the students who originally wrote the selection tests. The study does not include those who wrote the selection tests but were not accepted into the MBA degree. It also does not include those who were accepted into the MBA degree but did not complete the academic requirement in terms of completing the ten principle and core courses on which the MBA degree is based.

5.7.1. Characteristics of the Sample

After structuring and data cleaning the sample consisted of data from 741 MBA full and part time students who wrote the WBS's Admissions Test between 1998 and 2004 and had completed their MBA principle and core courses. 429 (57.9%) MBA students in this sample had completed the academic component of their MBA studies while the remaining 312 (42.11%) had

not completed the degree, for the reason that they had not handed in their research reports. The latter group made up the non-successful/failure category.

Those students who wrote the General Management Admission Test were not included in this study. The reason for this is that the researcher's interest in this study lies on the predictive validity of the four NIPR tests which were instruments developed in South Africa during the apartheid era but have had continued use in the years post the advent of democracy in 1994.

5.8. Procedure for Data Analysis

The literature on research design suggests that issues pertaining to data analysis should be carefully considered when designing a study, since the aim of data analysis is to transform information into an answer to the original research question (Bordens & Abbott, 2008; Rosenthal & Rosnow, 1991). The procedures used for data analysis in this study involved use of descriptive, correlational and comparative statistics.

5.8.1. Descriptive Statistics

Variable	Summary Statistics
Gender	Male=543 Female=198
Average Age	Males=30.96 (31) Female=31.2 (31)
Race	Asian=110 (14.8%): mean age=29 Black=182 (24.56%): mean age=31 White=449 (60.59%): mean age=31

Table 3: Descriptive Statistics of the Demographic Variables

According to the summary statistics displayed in table 3, WBS MBA students represented different groups (racial, gender, language and socioeconomic status) within the South African context.

5.8.1.1. Gender Distributions

Of this attenuated sample of 741 students, 543 were males and 198 were females. According to Kagan (2003), 28% of the South African MBA class is women. The WBS stood at 26.7%, which is comparable to the British and European findings (Kagan, 2003).

The composition of the South African MBA class has been dominated by English speaking males (Doody, 1997; Kagan, 2003). This classroom situation mirrored the South African economy. (Ramodibe, 2008). According to the Department of Labour's commissioned employment equity report, of the many employees recruited for senior management positions, 44.7% were White males, while 16.7% of that group were Black males (Ramodibe, 2008). Since the requirements for an MBA is some management experience this explains why the majority of participants are White male students. In most South African organisations White males dominate management and professional positions (Carmichael, 2003; Holburn, 1991b). However, the South African industry is seeing the increasing demand from Black students. J. M. Taylor 1977 in Erwee (1981) had already noted this change. He said that as Blacks become more educated, they are moving into more demanding positions in commerce and industry. However, the movement into industry has been slow.

Besides gender disparities in the data set, there was a small number of Coloured students, and for this reason Black and Coloured students were combined to form one Black (as opposed to White) category in the data set. The reason for doing so was that it was felt that having a

disproportionately small number of Coloured students in the study (approximately 28), would introduce complications into the analysis. Similarly, as the numbers of Chinese students were relatively small (12), the Indian and Chinese students were combined to form a single Asian category in the data set.

5.8.1.2. Age Distributions

Compared to the USA's 24-29 age range, the South African MBA class is more mature, and comes with solid work and managerial experience (Kagan, 2003). The South African MBA mean age is 33 with the range lying between 27-47 (Kagan, 2003). The average age of the sample from the Wits MBA programme was 31 years. This still aligned with the South African trends.

5.8.1.3. Qualifications Distributions

According to Kagan (2003), MBA programmes attract students from a wide range of academic backgrounds including Arts, Humanities, Law, Commerce, Engineering, Natural Sciences, Education, Computer Science, Medicine, Agriculture and the Military. She further mentioned that the diversity of students in the MBA class, gives students an "insight into employment opportunities that they might not previously have considered "(p. 5). In addition, the discussions in the classroom become diverse because students would refer to examples based on what they know, which most often will be in their fields of expertise.

The sample of students from the WBS sample confirmed Kagan's finding discussed above. The lowest qualification for the group was matriculation, with the highest qualification being the Doctor of Philosophy degree (PhD) or Doctor of Medicine. Most candidates in the WBS MBA class came from the Commerce, Engineering and Science background, in that particular order.

In terms of completion rate, these are the candidates with the high likelihood of completing their MBA studies.

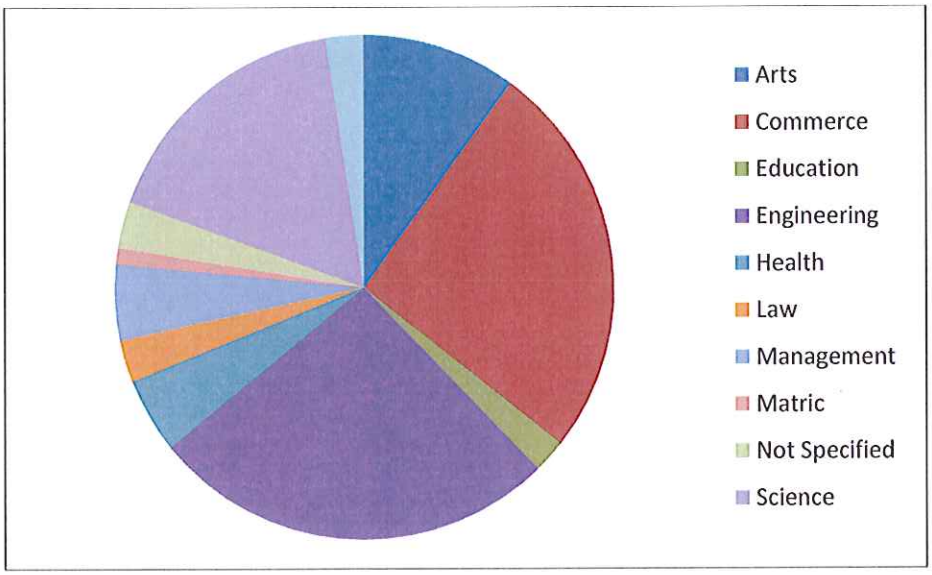


Figure 2: Distribution of Qualifications of WBS MBA Students

5.8.1.4. Language Distributions

Another aim of this study was to establish the differences between the different language groups. The language information from the assessment data reflected questions in the admission form concerning whether the candidates' language of preference for communication purposes, and not whether English was their home language. The data thus reflect that the majority of students opted for English (which is the medium of instruction at the institution - but might not be their mother tongue). It does not reflect whether English was the student's first or home language.

It is thus necessary to interpret the analysis relating to the language variable with caution. Approximately 273 White students preferred English, 98 Asian students mentioned that they

preferred English compared to 49 Black students who stated they preferred English as a language for communication. In contrast, 112 White students elected to communicate in a non-English language (quite often Afrikaans). However, there were also students who mentioned Greek, Spanish and Yugoslavian as their language of preference.

These factors are mentioned at this point, for the reason that the diversity in the language data makes it difficult, if not impossible, to accurately determine the language differences within the sample as a whole. What was clear from the data is that the majority of students selected English as their language even though it was not their mother tongue. This could be coded as a variable for comparative purposes. However, it was not possible to code additional distinctions within the language variable. This should be borne in mind as a limitation of this study.

5.8.2. Summary of Descriptive Statistics

Race	Completed Studies	% of total	Failed to Complete	% to total	Total	%
Asian	59	13.75%	51	16.35%	110	14.84 %
Black	86	20.05%	96	30.77%	182	24.56 %
White	284	66.2%	165	52.88%	449	60.59 %
Total	429	100%	312	100	741	

Table 4: Racial Distributions of those who were Successful versus those who were Not Successful to Complete their MBA Studies

Descriptive Statistics						
Variable	N	Mean	Std Dev	Median	Minimum	Maximum
Mental Alertness	741	26.42645	6.52451	27	4	41
Reading Comprehension Test	741	15.95547	4.03122	16	0	26
Arithmetic Reasoning Test	741	17.87854	6.64455	18	0	30
Conceptual Reasoning Test	741	24.46019	5.86078	25	4	35
DDM/Analytical Methods	741	64.8502	9.49441	65	38	87
Finance/Financial Management	741	63.34953	9.01264	64	38	90
Industrial Relations	741	65.85155	5.85774	66	47	85
Information Knowledge Management	741	67.63563	6.18735	68	42	88
Management Accountant	741	66.80027	6.94986	67	42	84
Marketing Management	741	66.61404	5.73274	67	34	82
Operations Technology	741	66.62213	5.59249	67	49	96
Organisational Design and Development	741	66.48583	7.08568	66	34	87
Strategy Management	741	67.30094	5.55536	68	35	84
Research Method	741	65.36572	7.79118	66	34	87
Average Core Courses	741	66.08758	4.20141	66.1	55.3	79

Table 5: Summary Statistics of the Scores for the Independent Variables

The descriptive statistics presented in table 5 above shows the measures of centrality for each independent variable and also how widely distributed the scores of the variables were. According to the data above, there were students who found the Reading Comprehension test and Arithmetic Reasoning test difficult, as the minimum scores for these tests were 0. Though

the tests might have been attempted, these students were unable to obtain one correct answer. These results will be further supported by data from the analysis later in the report.

According to these summary statistics, mean scores for the MBA core courses ranged between 63 and 67. Finance was the lowest with a mean of 63.3, followed by Data and Decision Making at 64. Both courses are popularly known as difficult courses, with high failure rates (Z.S. Koyana, personal communication, June 30, 2010). Courses like Strategic Management, Organisational Design and Development, Marketing Management and Research Methodology also featured among the courses with lower minimum scores, of less than 40. One would argue that these are based on the level of difficulty of the courses, but it could be as a result of many factors. Some of them being that a few students found the courses difficult, or because of circumstances they did not submit or write exams and hence negatively affected their overall scores in the courses.

The standard deviations from Data and Decision Making and Finance were higher (9.01 and 9.49 respectively) than the other courses, indicating that there was wide variability within the marks obtained by MBA students in these two courses. Other MBA courses did not show similarly high variability in their scores as standard deviation scores were below 7. With this sample of 741 participants, the researcher aimed to answer the following questions, to assist the WBS to determine the relevance of the tests in their selection process.

5.8.3. Inferential Statistics

The procedures used for analysing these data were as follows:

5.8.3.1. Research Question 1

- Are there significant relationships between the four NIPR tests and academic performance in the WBS MBA curriculum?

- Do the first order correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?

Independent Variables	**Scales of Measurement**
Mental Alertness Test scores	**Interval**
Reading Comprehension Test scores	**Interval**
Arithmetic Reasoning Test scores	**Interval**
Conceptual Reasoning Test scores	**Interval**
Dependent variable	
Academic Performance (based on average of ten core courses)	**Interval**

In order to investigate the relationship between variables, as well as to establish both strength and direction of the association, Pearson's Product Moment Correlation Coefficient was used. This statistic is normally used in studies in which the aim is to evaluate the direction and degree of the relationship between scores (Bordens & Abbott, 2008; Howell, 2008; Kachigan, 2001). The Correlation Coefficient indicates the degree of the relationship (e.g., 0 means no relationship), and the sign indicates the direction of the relationship (Bordens & Abbott, 2008;

Foxcroft & Roodt, 2001; Howell, 2008; Kachigan, 2001; Rosenthal & Rosnow, 1991). Positive correlation indicates a direct relationship where the increase in one variable is associated with an increase in the other variable. A negative correlation indicates an inverse relationship where an increase in one variable is associated with the decrease in the next variable.

By asking this question, the researcher aimed to establish the first order correlations that exist between the tests, the core courses, the average of the ten core courses as an interval outcome variable, and also the successful/non-successful completion as a dichotomous outcome variable. Another reason was to check for very high intercorrelations between the four NIPR tests, ten core courses and academic performance which might indicate the existence of multicollinearity, or a state of very high intercorrelations or inter-associations among the independent variables (Bordens & Abbott, 2008; Breytenbach, 2008; Howell, 2008; Kachigan, 2001; Rosenthal & Rosnow, 1991).

Multicollinearity can be regarded as an area of potential disturbance in the data, and if present, the statistical inferences made about the data may not be reliable (Galpin, 2001; Howell, 2008). Highly correlated variables contribute redundant information as quite often they measure the same thing (Dockrat, 2005). Thus, the first step is to establish whether there is a high likelihood that multicollinearity exists. If this is likely, one possible solution is to eliminate one of the highly correlated variables from the analysis on the assumption that if one does this, little would be lost if this were to happen. Another would be to undertake analysis of the data before inputting it into the regression model, and eliminating those observations which are outliers in the distribution (Bordens & Abbott, 2008). Another would be transforming the data in such a way that the distribution curve is smoothed. Another would be to undertake particular types of multiple correlation which would take account of (and essentially control for) the existence of the very multicollinearity or high inter-relationships existing in particular areas of the data (Bordens &

Abbott, 2008), and then to conduct a post hoc analysis on the residuals to see if there are particular observations which have actually distorted the results obtained.

In addition to answering sub questions relating to multicollinearity, through Research Question 1 the researcher was able to address a number of sub-questions concerning the first order relationships between the four NIPR tests and the individual academic performance variables in the predictive model.

1a) Is there a significant relationship between the Mental Alertness Test, Reading Comprehension Test, Arithmetic Reasoning Test, Conceptual Reasoning Test and academic performance as defined by the average of the ten core courses?

1b) Is there a significant relationship between Mental Alertness, Reading Comprehension, Arithmetic Reasoning, Conceptual Reasoning Test and the ten core courses?

1c) Is there a significant relationship between the ten core courses? What conclusions can we draw from such a relationship?

5.8.3.2. Research Question 2

- Do the multiple correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?

- Do the four NIPR tests predict academic success in the WBS MBA programme?

Independent Variables	Scales of Measurement
Mental Alertness test scores	**Interval**
Reading Comprehension Test scores	**Interval**
Arithmetic Reasoning Test scores	**Interval**
Conceptual Reasoning Test scores	**Interval**
Dependent variables	
Academic Performance (based on average of ten core courses)	**Interval**
Academic Success (Successful completion or Failure to complete the MBA	**Nominal**

As the criterion data were both interval and nominal, this question was twofold and required two different types of statistical analysis. The predictive relationship between the four NIPR tests in combination with the academic performance criteria was first analysed using forward stepwise regression. Post hoc analysis of the results was then undertaken to determine whether the predictive model could be accepted as accurate with a high degree of confidence.

The second part of the analysis then involved examination of the relationship between the four NIPR tests in combination and the nominal dependent variable of academic success. This was undertaken using logistic regression.

5.8.3.2.1. Assumptions of Multiple Regression

Multiple Regression is often undertaken for predictive purposes, to determine the relationship between a set of predictor variables (Independent Variables) and a single criterion variable (Dependent Variable) (Bordens & Abbott, 2008). The statistical analysis normally comprises of

two parts. First, a correlation matrix is developed, reflecting the first order relationship between all variables to be included in the regression model.

Multiple regression then follows, based on the model in which the predictive variables are examined. The strength of the relationship in each case is then estimated using the r^2 value, which represents the proportion of variance in the dependent variable that is explained by the variance in the Independent Variable (Howell, 2008).

In order for multiple regression to yield undistorted results (Bordens & Abbott, 2008), there is an assumption that:

- The data are normally distributed,
- Outliers in the data do not exert disproportionate influence, so as to distort the results,
- The first order relationships between the variables are linear ones and
- Multicollinearity does not exist in areas of the data set, so as to exert disproportionate influence and distort the results.

These checks were undertaken prior to as well as post hoc in the analysis, and will be reported in the results chapter.

5.8.3.2.2. Stepwise Multiple Regression

In Stepwise Multiple Regression, variables are entered into the equation according to the one accounting for the most variance in the dependent variable (Bordens & Abbott, 2008; Galpin,

2001; Kachigan, 2001). Variables are then entered one at a time until none of the remaining variables add significance to the regression coefficient.

Stepwise Multiple regression is used as a screening procedure to identify the most important and influential subset of independent variables (Dockrat, 2005; Galpin, 2001). Forward stepwise regression was identified as the best approach in the study, because with this approach, all simple linear regression models are considered in order to identify the one that gives the best regression (Galpin, 2001; Kachigan, 2001). The variable with the best prediction (or most significant) is added first, and then others follow in the stepwise procedure, depending on how much they add to the equation. (Breytenbach, 2008; Dockrat, 2005; Kachigan, 2001). The process stops only when all coefficients are significant at the desired level (Galpin, 2001).

Galpin (2001) highlights a number of potential drawbacks with this procedure, saying that it is not able to cope with suppressor variables. These are variables that are important together, not individually. If it is suspected to be the case, stepwise regression can be used. With backward elimination, for example, all variables are brought into the regression equation (Galpin, 2001). The coefficient of each variable is then examined in turn, and the variable with the least significant coefficient is removed (provided that the significance is above the threshold level of 5%). The regression equation is then recalculated, and the process is repeated. The process stops when all regression coefficients, are significant at the desired level.

In this study, forward stepwise regression was used, as it suited the researcher's purpose to know the variable predicting the highest first, as opposed to in reverse order. Through forward stepwise multiple regression, the researcher attempted to determine the degree to which predictive variables were able to predict the results (academic performance). This procedure was appropriate for the interval dependent variables.

However, for the nominal dependent variable (academic success), a different approach was necessary (Bordens & Abbott, 2008). Hence, logistic regression was used.

5.8.3.2.3. Logistic Regression

Logistic regression is essentially a regression model that is tailored to fit a categorical dependent variable (Allison, 2006). It allows one to predict a discrete outcome, such as group membership, from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these (Breytenbach, 2008; Peng, Lee, & Ingersoll, 2002). Generally, the dependent or response variable is dichotomous, such as presence/absence or success/failure. Discriminant analysis is also used to predict group membership with only two groups (Bordens & Abbott, 2008; Galpin, 2001; Kachigan, 2001). The reason for the use of logistic regression instead of discriminant analysis in this study was because the procedure had fewer assumptions. This provided the researcher with more room to run the statistics without concerns about the ability of data to meet certain requirements.

Logistic regression was also used as the literature indicated that the procedure was most suited to situations in which a researcher is trying to predict whether something "happens" or not (Breytenbach, 2008; Peng et al., 2002). In this case, the researcher was interested in whether or not the students had graduated or not. While discriminant analysis could have been used to answer the same question, it would have been from a different standpoint, of trying to establish whether the four NIPR tests could predict group membership and the probability of this occurring.

Bordens and Abbott (2008) suggest that logistic repression is particularly useful where the dataset is very large, and the predictor variables do not behave in orderly ways, or obey the

assumptions required of discriminant analysis (Bordens & Abbott, 2008; Galpin, 2001; Kachigan, 2001). The data set in this study was very large (741). Inspection of the data also indicated the presence of outliers, which would need to be accounted for in the analysis. Logistic regression was thus chosen as the less complicated of the two procedures.

5.8.3.3. Research Question 3

- Do the four NIPR tests predict academic performance equally for the different racial and language groups enrolled in the WBS MBA programme?

- Do the four NIPR tests predict academic success equally for the different racial and language groups enrolled in the WBS MBA programme?

Independent Variables	Scale of Measurement
Mental Alertness Tests	**Interval**
Reading Comprehension Test	**Interval**
Arithmetic Reasoning Test	**Interval**
Conceptual Reasoning Test	**Interval**
Dependent Variables	
Academic Performance(Average of the ten core courses) Academic Success (defined by the successful and failure to complete the MBA studies)	**Interval** **Nominal**
Classification Variables	
Race	**Nominal**
Language	**Nominal**

In order to answer question three, it was first necessary to run stepwise multiple regression for the interval dependent variable using the demographic variables (race and preferred language) as classification variables. The nominal dependent variable was analysed using Logistic Regression.

However, it was also necessary to compare the means of these groups, to confirm the existence or non-existence of group differences. This was undertaken with Analysis of Variance (ANOVA) and t-test statistics.

In order to confidently conduct Anova and t-test, it was important to check for equality of variance. The studentized residual plots were used as a means of determining equality of variance. Additive means was assumed when other assumptions had been met.

T-test was an option for determining group differences between the language groups since t-test works well with independent variables that have two levels (Bordens & Abbott, 2008). T-tests for two independent samples was the option for the current study as these focus on comparing means of samples that are independent from each another. ANOVA was then used an appropriate procedure for comparing the variance between the three racial groups.

5.9. Instruments

The archival data for this research were obtained from two sources, the assessment tests and performance on the MBA programme. In order to answer the research questions, the research design for this study examined the relationship between data from these two sources, examining scores from the NIPR tests, academic performance (defined as the average of marks obtained by the students in their ten core course) and academic success (defined as the successful completion of the MBA degree).

The NIPR tests used will not be discussed any further as this had already been done in chapter 4. Suffice to say that the WBS admissions office used scores yielded by the battery of selection tests which included the Mental Alertness Test, Reading Comprehension Test, Arithmetic Reasoning Test and Conceptual Reasoning Test. Biographical data from the student's admission form and previous academic record was then also considered in the admissions decision.

The data relating to the dependent variables for the study were obtained from the student's academic performance records. The data relating to academic success were also obtained from the same source (the MBA academic performance record for each student). The average of the students' marks in their ten core courses were used as an index of academic performance, while academic success was determined by a dichotomous completion or non-completion status for the degree as a whole.

One issue in considering data from a dichotomous variable is whether there is likely to be sufficient variance in the data from both categories in the dichotomous variable to yield undistorted results. In this study, the lack of completion of the MBA degree (completion rate) was surprisingly low. This effectively meant that the ratio between those who completed the degree versus those who did not was sufficiently balanced to not skew a dichotomous analysis.

However, the granting of exemptions from principle courses resulted in gaps within the data set. Some students did not have marks for the principle courses, because they could have previously completed a management course with the school, or they held undergraduate qualifications in a similar course. This presented a potential problem of missing values, which was not possible to resolve. As those student records which indicated that students had not completed the full academic component included students with exemptions as well as those students who had dropped out of the degree, and as there was no means of checking from the available data which student fell into which of these two categories, the decision was taken to exclude all student records in whom there were missing data from the analysis. This has already been discussed as a limitation applying to the data available for analysis in this study.

5.9.1. Descriptive Data on the Four NIPR Tests

5.9.1.1. The Distribution of Scores on the Mental Alertness Test

To be successful on the Masters of Business Administration (MBA) students requires a high level of reasoning ability. Students are given a mark for each correct answer chosen from the multiple choice questions. The scores are then added to determine the raw score for the test.

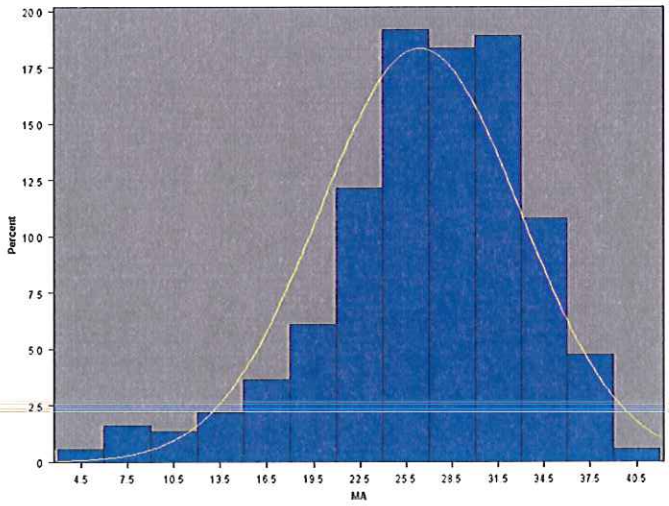


Figure 3: Normal Distribution Graph for the Mental Alertness Test

According to the histogram presented above, the Mental Alertness test scores were widely distributed with ranges between 4.5 and 40.5. Most scores lie between 22 and 34.5. Kurtosis is 0.928 and skewness is -0.8484. According to Breytenbach (2008) kurtosis is problematic when it exceeds 7. Similarly, skewness becomes a problem in regression when it exceeds 2. None of the distributions for the four NIPR tests had values that exceeded these suggested values.

5.9.1.2. The Distribution of Scores on the Reading Comprehension Test

The English Reading comprehension test, tests the ability to understand written English. The graph below looks almost symmetrical, as the values lie between 6 and 26 and show a bell shaped graph. Kurtosis was 0.949 and skewness at -0.44.

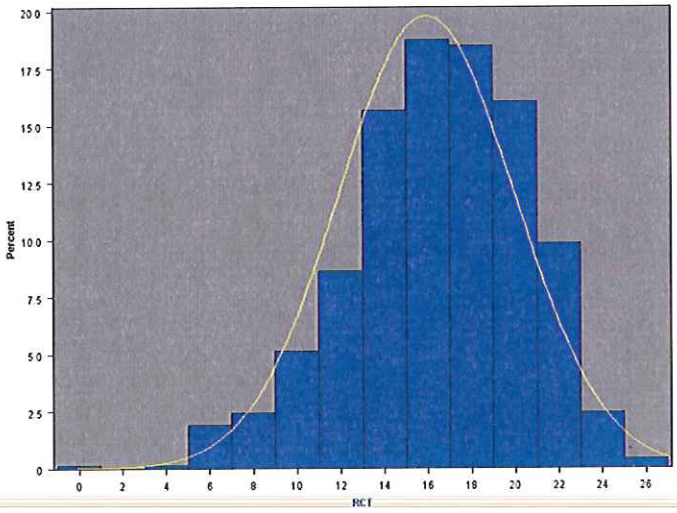


Figure 4: Normal Distribution Graph for the Reading Comprehension Test

However, there were outliers pulling the graph to the left. The steps taken to establish the influence of these on the analysis will be referred to in the next chapter.

5.9.1.3. The Distribution of Scores on the Arithmetic Reasoning Test (Standard Level)

This test measures understanding of the fundamental rules of arithmetic and application of these rules in logical strategies in order to solve problems of a purely numerical nature (J. M. Taylor & Halstead, 1985 in Halstead 1985). It is useful in predicting success in activities involving the manipulation of numbers or symbols representing numbers.

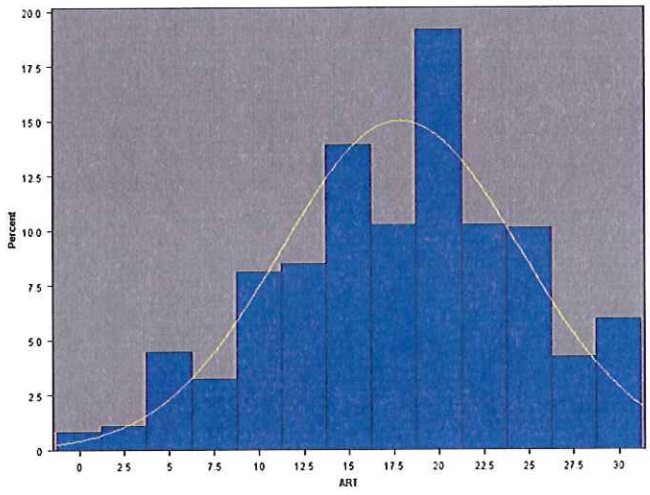


Figure 5: Normal Distribution Graph for the Arithmetic Reasoning Test

The distribution for Arithmetic Reasoning appeared to be normal. There was a reasonable number of those who struggled with the arithmetic reasoning test, as well as a number of students who found the test relatively easy. However, unlike the two histograms presented above, the Arithmetic Reasoning graph had a lower peak with kurtosis equal to -0.388 and skewness at -0.291. There was some skewness to the negative side of the distribution. However, even in the two tails of the curve the data still clustered together, indicating that despite some very low and very high scores, the shape of the distribution broadly followed a normal distribution.

Unlike in the distribution of the Reading Comprehension, the distribution of the Arithmetic Reasoning Test did not indicate real outliers in the data. Nevertheless the potential of outliers to influence the analysis was considered and will be reported on in the next chapter.

5.9.1.4. The Distribution of Scores on the Conceptual Reasoning Test

This is a non-verbal measure (or performance test) of inductive reasoning ability. The scores from the graph below are skewed to the left. Scores range between 5 and 35, with most scores clustered around 20 and 30. The distribution is pulled to the left by the extremely low scores ranging between 4 and 10. Kurtosis for Conceptual Reasoning test was 0. 3116 and skewness was -0.711. Clearly negative skewness existed but it was not big enough to be a concern (being below 2).

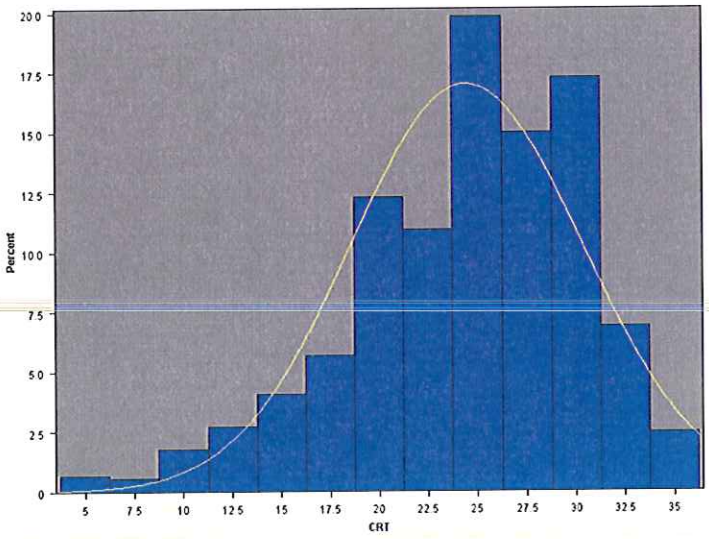


Figure 6: Normal Distribution Graph for the Conceptual Reasoning Test

However, the graph did indicate the presence of certain low scores pulling the graph to the left. The steps taken to establish the influence of these on the analysis will be referred to in the next chapter.

5.10. Overall Trends in the Distributions of Scores on the Independent Variables

Overall, the distributions of scores for the four NIPR tests tended to be negatively skewed. This implied that more students gained higher scores than lower scores on these tests. There were some outliers in the data, indicating that this factor would need to be taken into account in the analysis. Nevertheless, given the sheer size of the data set and the indications that kurtosis was below the level considered critical in influencing a multivariate analysis (Bordens & Abbott, 2008) it was considered unnecessary to transform the data prior to analysis.

5.11. Academic Performance and Academic Success

Often academic performance and academic success are used interchangeably and the meaning of each term gets lost in those translations. However, the two terms are different and it is important to clarify this. Academic performance may vary from institution to institution, but for most institutions it is when a student obtains an average final mark of 50% for a particular course (Riordaan, 2002). In Mokone's study, academic performance is defined as grades obtained by students in their respective engineering courses (Mokone, 2008, p. 7).

Academic success on the other hand is defined by the successful completion of the actual degree one is registered for (Fraser & Killen, 2003; MacFarlane, 2006). When a qualification has been accomplished, it is then that we talk of academic success. That implies a state of achievement.

In the Wits Business School, the Faculty Office relies on students' performance on the six MBA fundamental courses, ten core courses, five electives and a Research Report to determine academic success. The individual mark (a syndicate mark in some courses) and the

examination are the components of a final mark for each of the 21 courses that students study in the MBA programme. The averages of both are added to compile the final mark.

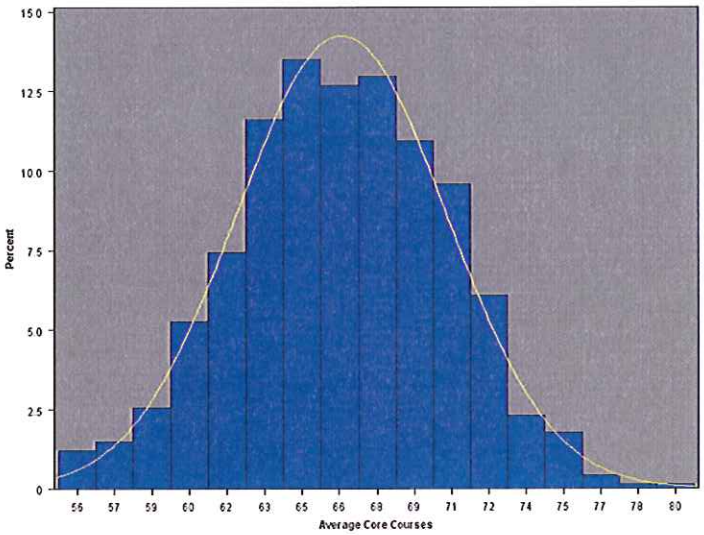


Figure 7: Normal Distribution Graph for the Scores of the Average of the Core Courses

The scores for academic performance show a somewhat normal distribution, with 2 or 3 extremely high values that drag the distribution to the extreme positive direction. Most values lie between 60 and 72, meaning the performance of the majority of students was in C (60-69) category. The mean for average of the ten core courses rests at 66.09 with Standard Deviation of 4.2.

As with the predictor variables, the shape of the distribution implied that it was likely, given the size of the data set, that the data were normally distributed enough to be submitted to multivariate analysis (Bordens & Abbott, 2008).

5.12. Procedure

The research began after the University's Ethics board had granted ethical clearance and the researcher had obtained permission from the Head of School at the WBS, Head of Academic Programmes at the WBS and the University's Registrar. Seeing that the majority of participants had already completed their MBA studies, the University Registrar acted on the behalf of students.

As indicated in previous chapters, the data were historical data which had been archived. They were thus available data which had been collected in two phases:

Phase 1:

In the process of applying for the Master of Business Administration course, candidates wrote the tests as part of the selection process. The test scores were then scrutinised with other relevant biographical information which the students submitted with their applications. Based on this scrutiny of pencil and paper information, acceptance and non-acceptance into the programme took place.

Phase 2:

The second phase involved collating of the candidates' performance on the various courses they took as part of the MBA programme from the first year until the last year, when they submitted their research report. The student record was then stored electronically.

As these two phases were linked, the researcher was able to collect and collate the testing information with information from the faculty office in a single spread sheet. To ensure that the students' anonymity was secured, once the students information has been saved, names and student numbers were deleted and the students were numbered arbitrarily. This step was taken to make sure that no one could link information from the analysis back to any particular student. This was also done to ensure confidentiality of students' private information. Once these steps in data collection had been completed, the data were analysed with the statistical procedures as discussed in the earlier sections of this chapter.

5.13. Summary

Being based on available data drawn from the university's archives, a non-random (purposive) sampling strategy was used in obtaining the sample of data for this study. The researcher relied on data that already existed. Hence, the sample was non-random, representing data from the records of Wits MBA students.

The research design described in this chapter was non-experimental. In order to answer the research questions, correlational research was undertaken. The statistical procedures used involved an initial step in which first order correlations were calculated. This was followed by multivariate analysis using stepwise multiple regression and logistic regression.

There are a number of limitations applying to the data set. The first was that the data analysed represents an attenuated sample. The Wits archives contained data relating to +/- 2000 MBA students. However, the majority of these student records had missing data, and there was no way in which the problem of which students had exemptions and which students had dropped out of the programme could be addressed as the original student files were no longer

accessible. The decision was thus taken to work only with those student records which were complete. This implied attenuating the sample, and not dealing with those student records relating to students who had dropped out of the MBA programme, or those student records which were incomplete for other possible reasons including exemption from course requirements.

The sample with which this study deals is thus severely attenuated, and represents an attenuated sample of 741 students. This is a limitation of this study. Nevertheless, it is suggested that, being attenuated, the data analysed in this study is potentially more internally consistent, consisting of the records of those students who satisfactorily completed the academic component of the MBA degree. This sample consists of 429 student records representing the academic performance of those students who completed the academic and the research requirements of the degree and graduated, and 312 students who did not.

The research design described in this chapter relies on first order correlation to answer the question about the relationship between the four NIPR tests, ten core courses and academic performance (as defined by the average of the ten core courses comprising the MBA degree). The research design for the other research questions also relies on multivariate statistics. Forward Stepwise Multiple Regression has been used as a procedure for establishing the degree of overlap between the four NIPR tests in predicting academic performance, and Logistic Regression has been used to answer the same question relating to the dichotomous dependent variable of academic success in either graduating in the MBA degree or not.

The last research question focuses on whether the four NIPR tests predicted academic performance and academic success differently for students in the sample from different

language and racial groups. This question was addressed using Forward Stepwise Multiple Regression and Logistic regression, ANOVA, and t-tests.

In operationalizing a study, the appropriate use of statistics depends on assumptions concerning the use of these particular statistical procedures with quantitative research data. These assumptions have been described in this chapter, together with the steps taken to ensure that this particular data set is suited to the univariate and multivariate procedures used in the analysis. The next chapter presents the results from the study, relative to these assumptions.

6. Results and Interpretation

6.1. Introduction

The previous chapter has outlined the assumptions of the multivariate statistical procedures conceptualised as appropriate to answering the research questions on which the research design is based. To answer the research questions, a number of statistical analyses were carried out using SAS, with 0.05 set as the alpha level which would be adopted as a level of significance appropriate to answering the research questions pertaining to the study. Levels of significance beyond this level (e.g., $p < 0.001$) would then also be reported.

In reporting the results of the study, a number of tests conducted to establish whether the assumptions applying to the statistical procedures used in this study will also be reported. This implies some overlap between the narrative in the previous research design chapter and the current chapter, in which the results of the research are reported.

6.2. Assumptions of Parametric Measures

There are a number of assumptions relating to use of parametric statistics (Bordens & Abbott, 2008; Galpin, 2001; Howell, 2008; Kachigan, 1991). These can be summarised as follows:

Random Independent Samples: Each individual in the population has an equal chance of being selected for inclusion into the sample. In this study, sample depended on the availability of data. No scientific approach was utilised in selecting the sample. Though, ideal, it was not possible to meet this assumption.

As a result of a non-random sample, generalizability cannot be made with any level of confidence. This should thus be noted as a limitation applying to the results of the research reported in this chapter.

Interval Scale: The data from both the four NIPR tests forming the independent variables and the academic performance average forming the dependent variable in this study are at an interval scale of measurement, in that the spacing between values along the scale is known, and zero does not indicate the absence of a quantity (Bordens & Abbott, 2008). "With interval scales, one is able to determine that one unit is larger or smaller than another and by how much" (Bordens & Abbott, 2008, p. 137). This is due to the fact that equal differences between scale values have equal meaning (Kachigan, 1991).

This was the case with the data from the four NIPR tests and the core courses. However, the race and language variables are nominal. So was the success/fail criterion outcome. Different statistics were thus necessary to conduct predictive analyses with nominal data.

Normal Distribution: To examine the distributions, histograms and the Kolmogorov Smirnoff test of normality were used to determine whether the data were normally distributed. In the first instance, the researcher relied on pictorial representation of the data through histograms. These graphs produced data that depicted a normal distribution with some minor skewness and have been presented in the previous chapter (See Appendix B).

If one was relying solely on these histograms, with a large sample it would be reasonably safe to conclude that a sufficiently symmetrical distribution of variables existed to not compromise or distort the analyses. However, to be absolutely sure, Kolmogorov-Smirnov values were calculated, yielding the results in the table following.

Test		Statistics	P-value	
Mental Alertness	D	0.084185	Pr > D	<0.0100
Reading Comprehension	D	0.082005	Pr > D	<0.0100
Arithmetic Reasoning	D	0.063641	Pr > D	<0.0100
Conceptual Reasoning	D	0.089451	Pr > D	<0.0100

Table 6: Results from Kolmogorov Test of Normality

Overall, it can be concluded that as the Kolmogorov Smirnov revealed results that were not significant at $p<0.010$, less than a 1% likelihood of error stemming from an irregular (non-normal) distribution existed in using data from the four NIPR tests as independent variables in the analysis.

The decision to continue using parametric statistics with the sample of 741 was also taken (irrespective of the Kolmogorov-Smirnoff results), for the reason that the numbers in the data set were large enough. Any deviation from normality was thus felt to be unlikely to have a major effect on the results. Overall with a sample of 741 participants and with a number of independent variables involving psychometric tests with which certain of the student applicants might be unfamiliar, the reasoning was that it was likely that one would obtain minor deviations from normality affecting the analyses. It was thus important in addition, to establish whether extreme values would have, or actually had had, a disruptive effect on the multivariate analyses.

For this reason, the decision was taken to run outlier analyses. Outliers were defined as extreme scores that lay far from others in the distributions, indicating individual scores lying well outside the pattern of data (Moore & McCabe, 2006 in Bordens & Abbott, 2008). Sometimes the existence of these scores is not an error, but a true reflection of the data set. However, there is also the possibility that extreme values are due to errors in the data capturing process.

Box and Whisker Plots were drawn to determine the extreme nature of the outliers and their potential effects on normality both for the NIPR tests as well as for the individual courses on which the academic performance criterion was based (See Appendix C). What was found was that extreme scores existed in all the independent variable distributions with the exception of the Arithmetic Reasoning test, which was the only NIPR test which did not have extreme scores that could be classified as outliers. On the side of the dependent variable, Data and Decision Making was also the only core course forming part of the criterion variable of academic performance without outliers.

However, even though the Box and Whisker Plots revealed outliers for Mental Alertness, Reading Comprehension and Conceptual Reasoning test and the other core courses, the maximum number of these outliers was twelve, for Information and Knowledge Management. Considering the size of the sample, in consultation with the statistician advising on the study, the presence of twelve outliers was not considered to be a cause for concern.

Thus, normal distribution was confirmed both by formal testing and inspection of the distributions. This implied that based on these analyses, the researcher was able to confidently use parametric multivariate statistics that required normal distribution. In addition, the decision was taken to run post hoc analysis of the residuals to establish whether extreme scores had actually exerted an influence on the results (Galpin, 2001).

6.3. Results Relating to Research Question 1

- Are there significant relationships between the four NIPR tests and academic performance in the WBS MBA curriculum?

- Do the first order correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?

This question aimed to establish if there were relationships between the four NIPR tests and academic performance as defined by the average of the ten core courses. To answer the question required a twofold analysis.

6.3.1. Relationships between the NIPR Tests and Academic Performance

In the first analysis, a table of intercorrelations between the four NIPR tests and academic performance (as represented by the average on all ten core courses) was constructed. The results of this are presented in table 7 below. In the second analysis, intercorrelations were calculated between the four NIPR tests and all the individual courses in the MBA programme, as well as intercorrelations between the different individual courses in the MBA programme. Part of the full correlation matrix is presented later in this section, while the full table of intercorrelations is presented in the appendices as Appendix D (see the Pearson's Correlation Coefficients matrix).

	MA	RCT	ART	CRT	Academic Performance
MA	1	0.538 <0.0001	0.632 <0.0001	0.665 <0.0001	0.31 <0.0001
RCT		1	0.326 <0.0001	0.395 <0.0001	0.268 <0.0001
ART			1	0.567 <0.0001	0.23 <0.0001
CRT				1	0.27 <0.0001

Table 7: Results of Relationships between NIPR Tests and the Average of the Core Courses

Table 7 indicates that significant ($p<0.001$), yet moderately strong relationships were observed between the four NIPR tests and academic performance. Mental Alertness was the only test that showed a firm level of correlation with all other tests in the study, with the relationship between Conceptual Reasoning Test yielding a coefficient of 0.66 ($p<0.0001$). The score confirmed a strong relationship between the two variables, however the relationship was not in itself high enough to raise caution for multicollinearity. This was confirmed from the literature, in which a similarly strong correlation between Mental Alertness and the Conceptual Reasoning Test was reported by Boeyens (1990) in his study on White male matriculants at St Stithians College.

Correlations between Mental Alertness and Academic Performance produced a correlation coefficient of 0.32. The direction of the relationship was positive and of a moderately strong magnitude. This means that you can expect someone with high Mental Alertness test scores to also obtain relatively high scores in their MBA studies or low test scores associated with lower scores in the course. The relationships were significant ($p<0.001$), meaning there was a very low likelihood that these results are explained by chance.

Mental Alertness was the only NIPR test that showed a somewhat firm correlation with the dependent variable. Following it in terms of magnitude was Conceptual Reasoning Test at 0.27, Reading Comprehension at 0.268 and Arithmetic Reasoning at 0.23.

6.3.2. Relationships Between the NIPR Tests and the Ten Core Courses

The relationships between the four NIPR tests and the individual core courses in the MBA programme is reflected in the table below:

Core Course	MA	RCT	ART	CRT
DDM/Analytical Methods	0.29942	0.20418	0.26012	0.29585
	<.0001	<.0001	<.0001	<.0001
Finance/Financial Management	0.31775	0.19897	0.29393	0.27929
	<.0001	<.0001	<.0001	<.0001
Industrial Relations	0.06902	0.13558	0.03769	0.04895
	0.0604	0.0002	0.3055	0.1832
Information and Knowledge Management	0.10826	0.14682	0.04083	0.09697
	0.0032	<.0001	0.267	0.0083
Management Accounting	0.26983	0.25127	0.14292	0.15972
	<.0001	<.0001	<.0001	<.0001
Marketing Management	0.1004	0.05849	0.08302	0.06908
	0.0062	0.1116	0.0238	0.0602
Operations Technology	0.28889	0.1967	0.25409	0.30744
	<.0001	<.0001	<.0001	<.0001
Organisational Development and Design	0.11457	0.14132	0.03469	0.06503
	0.0018	0.0001	0.3456	0.0769
Strategic Management	0.1718	0.14765	0.09153	0.14894
	<.0001	<.0001	0.0127	<.0001
Research Methodology	0.09768	0.10516	0.05408	0.09376
	0.0078	0.0042	0.1413	0.0107

Table 8: Results of Relationships Between the NIPR Tests and the Individual MBA Core Courses

It will be evident from table 8 that a large number of significant relationships between the NIPR tests and the individual core course were found. Industrial Relations, Marketing Management and Organisational Development and Design, Strategic Management, Research Methodology were the core courses that had non-significant relationships with one or two of the NIPR tests. This was with the Reading Comprehension Test, the Arithmetic Reasoning Test and the Conceptual Reasoning Test.

Significant relationships were observed between Mental Alertness, Data and Decision Making, Finance/Financial Management, Information and Knowledge Management, Management Accounting, Marketing Management, Operations Technology, Organisational Development and Design, Strategic Management and Research Methodology. Overall, however, though Mental Alertness correlated at a significant level with the majority of the course marks, the majority of the relationships found were below r=0.30, and thus not firm.

Similar trends occurred in other areas of the data. In terms of the intercorrelations between the core courses (see the Pearson's Correlation Coefficients matrix in Appendix D), there were some firm correlations. Data and Decision Making, for example, correlated with Finance at r=0.52 (p<0.0001). There were also other firm correlations with Operations Technology at r=0.408 (p<0.0001). The presence of these relationships indicates that if an MBA student scores very well in the Data and Decision Making course there is a likelihood that they will pass Finance quite easily, and conversely that the more an MBA student finds Data and Decision Making difficult he or she is also likely to find Finance difficult. These were important indications which may be useful to those developing the MBA programme. Operations Technology was the one course that seemed to correlate well with almost all

other core courses, except Industrial Relations at r=0.20. It was also noteworthy that all correlations yielded significance (p<0.001). This was probably an artifact of the large size of the data sample.

The relationship between the four NIPR tests and the actual courses in the MBA programme revealed some significant and non-significant relationships (For further details see the Pearson's Correlation Coefficients matrix in Appendix D). However, the relationships were not firm, with the strongest at 0.317 between Mental Alertness and Finance and 0.31 between Conceptual Reasoning and Operations Technology.

Overall, in terms of the first research question, it is possible to conclude that a significant relationship exists between the four NIPR tests and academic performance as defined by the average of the ten core courses. The highest level of relationship was found between Mental Alertness and overall academic performance (r=0.31). However, for selection purposes this may not warrant continued use of the instrument, for the reason that a first order of correlation of 0.31 indicates that only 9% of the variance in the two variables is common.

There were also significant relationships with many other individual core courses taught in the MBA programme. The highest level of correlation was between Mental Alertness and Finance and Financial Management (r=0.32). Again, while this finding is interesting, it may not warrant continued use of the instrument for selection purposes, as this level of first order correlation also indicates that only 9% of the variance in the two variables is common. The majority of the intercorrelations between the core courses were less firm, even though with

the size of the sample, they showed significance.

Murphy and Davidshofer (2005, p. 191) have stated that in most instances, "When tests are used to make important decisions, they account for relatively small percentages of the variability in criteria". However the validity of a test as an aid in making decisions should be judged by evaluating the effects of tests on the accuracy of decisions using additional evidence.

In order to establish whether the NIPR tests yielded additional evidence when used in combination, regression analyses were conducted, which are reported in the following section.

6.4. Results Relating to Research Question 2

- Do the multiple correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?

- Do the four NIPR tests predict academic success in the WBS MBA programme?

As with the previous research question, Research Question Two required a twofold analysis. The dependent variable was defined as the average of the ten core courses and to establish the multivariate relationship of this dependent variable and the four NIPR tests, Stepwise Multiple Regression was used as the analysis tool. However, in predicting whether the sample of students were academically successful (i.e., whether they had been likely to successfully complete their MBA studies or not), Logistic regression was chosen as the most appropriate procedure to use with a dichotomous dependent variable.

The reason for this was that this statistical procedure works well with dichotomous dependent variables and there are fewer assumptions to meet than with a procedure such as Discriminant Analysis (Bordens & Abbott, 2008). As both multiple regression and logistic regression are multivariate procedures, there are a number of assumptions which need to be met before the statistics can be used and interpreted confidently. These are listed by Galpin (2001) and Bordens and Abbott (2008), and are summarized in the following section.

6.4.1. Assumptions of Multiple Regression

Sample size

Regression analysis requires at least 20 cases per independent variable in the analysis. In the simplest case of having just two independent variables that requires n > 40 (Howell, 2008; Osborne & Waters, 2002). This study had exceeded this requirement as there were 741 subjects for four independent variables.

Linear relationship between the independent and dependent variable(s)

Standard Multiple Regression can only accurately estimate the relationship between dependent and independent variables if the relationships are linear in nature (Galpin, 2001; Howell, 2008). If the relationship between independent variables (IV) and the dependent variable (DV) is not linear, the results of the regression analysis will underestimate the true relationship. Through the scatterplots, one can draw an imaginary line to determine the existence of a linear relationship (See Appendix E). An imaginery line can be drawn through the data, however there were few observations falling out of the graphs, and these would be identified as residuals.

Residuals are normally distributed

Another method of detecting nonlinear relationships is to examine residual plots (plots of the residuals as a function of standardized predicted values). Residuals are defined as the error of prediction (Howell, 2008). In order to determine whether residuals are normally distributed, it is important to look at the residual versus predicted value plot to see if the residuals are more or less symmetrically distributed about the zero line (zero on the vertical axis) (Galpin, 2001).

The figure below illustrates this type of distribution. Even though the fit was not perfect, it will be evident from figure 8 that the residuals were symmetrical to the normal distribution. Thus, the difference between actual and error of prediction scores was not big enough to cause concern about the possibility that residuals (error scores) had actually exerted a distorting influence on the results.

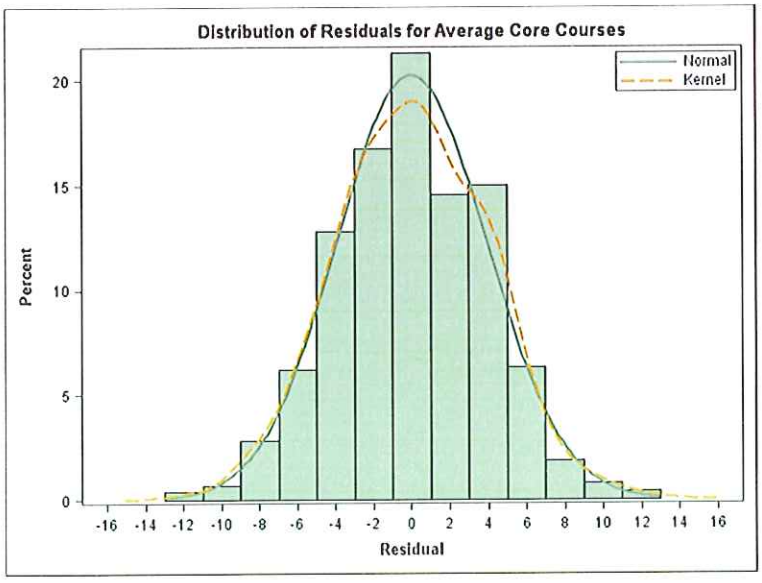


Figure 8: Distribution of Residuals for the Average of Core Courses

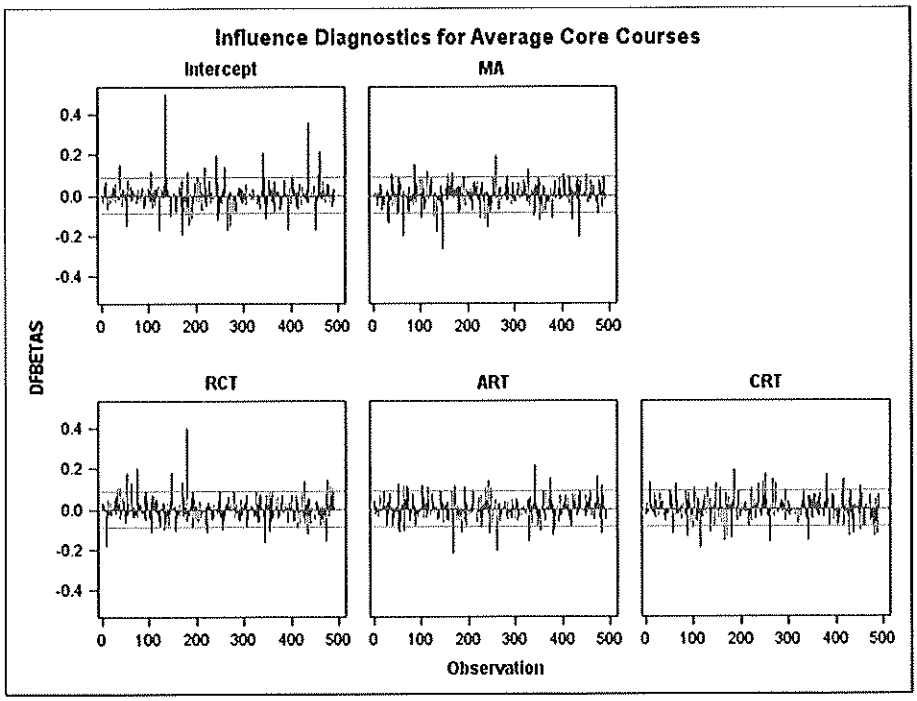


Figure 9: Results of Influence Diagnostics for the Average of Core Courses

The conclusions from the evidence presented in figure 9 and 10 was that the likely effect of the residuals was not major, either in the case of the independent variables or the dependent variable, for the reason that the majority of the residual values lay between +0.2 and -0.2. This can also be inferred from the effects observed in this study, as presented in figure 10 below.

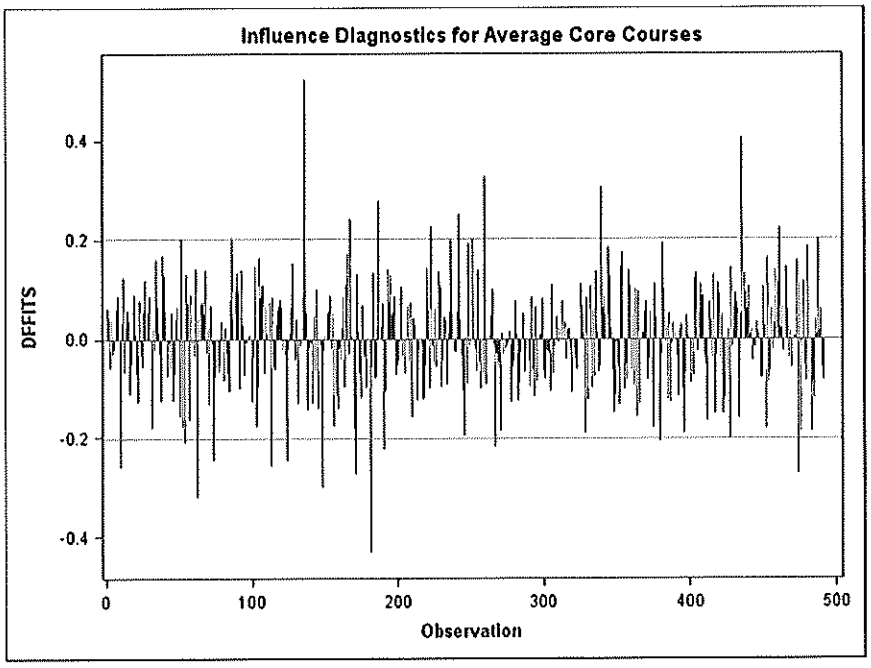


Figure 10: Results of Influence Diagnostics for the Average of Core Courses

Collinearity Diagnostics indicate that Multicollinearity is not a Problem in the Data

Regression is sensitive to multicollinearity. In order to make sure that the high correlation between Mental Alertness and Conceptual Reasoning Test would not have a negative impact on use of a multivariate procedure based on principles of regression, collinearity diagnostics were used. Collinearity diagnostics for the four NIPR tests are presented in table 9 below.

Collinearity Diagnostics							
Number	Eigenvalue	Condition Index	Proportion of Variation				
			Intercept	MA	RCT	ART	CRT
1	4.85113	1	0.00162	0.000945	0.00174	0.00274	0.00117
2	0.07375	8.11064	0.08768	0.000201	0.09991	0.60907	1.97E-05
3	0.03299	12.1265	0.27486	0.00167	0.7081	0.03821	0.16811
4	0.02555	13.77832	0.60136	0.15552	0.00178	0.30156	0.35162
5	0.01658	17.10486	0.03447	0.84167	0.18847	0.04843	0.47907

Table 9: Table of Results from Collinearity Diagnostics

In the above table, collinearity was checked using condition index tests and Eigen values. In cases where no collinearity existed, all Eigen values would be 1. Values smaller or larger than 1 would indicate departures from the ideal situation. Too small or too large Eigen values would indicate multicollinearity problems (Cohen and Cohen, 1975).

Ideally, a good condition index should be less than 10, indicating that the independent variables are not correlated. A condition index of 20 is also acceptable, but anything from 30 and above is not acceptable (Cohen & Cohen, 1975). Values above 30 would indicate that the independent variables were highly correlated, which would be problematic (Rosenthal & Rosnow, 1991). It is evident from table 9 that the condition indices for the four NIPR tests ranged between 8 and 17.

Thus all values fell below the 20 cut off point. Thus, though the matrix of first order correlations indicated that a significantly high correlation had been found between Mental Alertness and Conceptual Reasoning, this was not high to cause concerns that multicollinearity would distort the results obtained from use of a regression procedure.

All variables are in the equation

This assumption is normally tested by checking for quadratic or other trends in a residual plots check (Osborne & Waters, 2002). If the study focuses on five IV's, then all of them must be entered into the equation, for multiple regression to hold. $Y = ß_0 + ß_1X_1 + ß_2X_2 + ß_3X_3 + \ldots\ldots + ß_nX_n$, where Y indicates the dependent variable, $ß_0$ the constant of the equation, $ß_1$, $ß_2$, $ß_n$ the regression coefficients and X_1, $X_2 \ldots X_n$ the different independent variables. The residuals are then checked post hoc, to establish if there have been violations indicating that error distortion has occurred in the analysis. In this study, there were four independent variables, therefore the nth variable = $ß_4X_4$.

Errors have same variance throughout - Homoscedasticity

Homoscedasticity means that the variance of errors is the same across all levels of the IV (Galpin, 2001; Osborne & Waters, 2002). When the variance of errors differs at different values of the IV, heteroscedasticity is indicated. According to Berry and Feldman (1985, in Osborne & Waters, 2002), slight heteroscedasticity has little effect on significance tests. However, these authors argued that, when heteroscedasticity is marked, it can lead to serious distortion of results and seriously weaken the analysis, thus increasing the possibility of a Type I error.

In order to determine this, it is important to check residual plots for wedge, parabolic or other strange shapes, as well as to look at the plot against time (Galpin, 2001, Osborne & Waters, 2002) This assumption can be checked by visual examination of a plot of the standardized residuals (the errors) by the regression standardized predicted value.

Ideally, residuals are randomly scattered around 0 (the horizontal line) providing a relatively even distribution. Heteroscedasticity is indicated when the residuals are not evenly scattered around the line. The data distribution in this study indicated some homoscedasticity as per figure 11 below. In addition to the evidence already presented in Figure 10 above, it is evident from the graph that the majority of error scores for the dependent variable (average of the core courses) cluster evenly around the zero point.

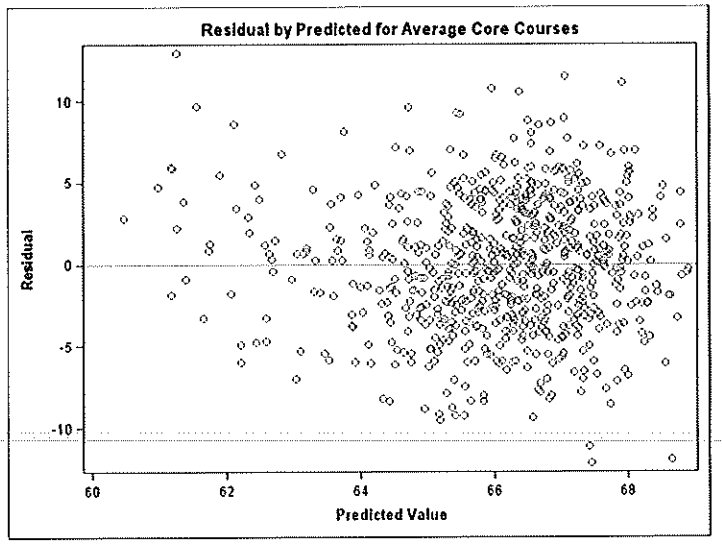


Figure 11: Residuals by Predicted for Average of Core Courses

Uncorrelated errors

Predictive studies tend to be affected by time lag between the collection of the criterion variable and the predictive variable. It is therefore important to check for sinusoidal patterns in the residual and time plots. The Durbin-Watson statistic can be used to look at the serial correlations (Galpin, 2001).

The Durbin-Watson statistic is used to detect the presence of autocorrelation (a relationship between values separated from each other by a given time lag) in the residuals from a regression (Galpin, 2001). Durbin-Watson statistic, ranges in value, from 0 to 4 with an ideal value of 2 indicating that errors are not correlated (although values from 1.75 to 2.25 may be considered acceptable) (Galpin, 2001). A value significantly below 2 indicates a positive correlation and a value significantly greater than 2 suggests negative correlation. An outcome of 2.055 was achieved in the current study, thus errors were not correlated.

6.4.1.1. Results from Stepwise Multiple Regression

The discussion in the section above has been undertaken to indicate that in this study, all the assumptions necessary for appropriate use of Multiple Regression as a statistical procedure were met, indicating that the data set was suited to this type of multivariate analysis. Having established this, it is possible to present and interpret the results from the various multivariate analyses with a high degree of confidence, as follows:

In this study, multiple regression did not confirm a firm predictive relationship between the four NIPR tests and academic performance (See Appendix F). According to table 10 below, overall use of the tests for selection purposes was only able to explain a small proportion (0.1201) of the variance in academic performance.

What this finding meant was that only 12.01% of the variability in academic performance can be accounted for by variability in the four NIPR tests together. This amount of predictable variability was not substantial, as information from the four NIPR tests together predicted less than one-eighth of the variance in academic performance. That implied that 88% of the variance in academic performance was unaccounted for, and could be explained by other unknown or

extraneous variables. This finding, in turn, indicated that it was very difficult to justify the continued use of this particular battery of tests for selection purposes in the WBS context.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	3	1568.652	522.8841	33.53	<.0001
Error	737	11494	15.59527		
Corrected Total	740	13062			
Root MSE	3.94908	**R-Square**	0.1201		
Dependent Mean	66.08758	**Adj R-Sq**	0.1165		
Coeff Var	5.97553				

Table 10: Results from Multiple Regression Analysis Between NIPR Tests and Academic Performance

Due to the large size of the sample of data examined, though the *F* value relating to the variance was significant (F=33.53; p<0.0001), the actual proportion of variance explained by the four NIPR tests was very small.

Further analysis was thus undertaken to establish which of the four NIPR tests was the best predictor of academic performance. This was done using a forward selection regression procedure which would take account of overlap between the four tests (Bordens & Abbott, 2008).

Summary of Forward Selection							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	MA	1	0.1013	0.1013	15.0643	83.33	<.0001
2	RCT	2	0.0131	0.1145	6.0797	10.94	0.001
3	CRT	3	0.0056	0.1201	3.3686	4.72	0.0302
Multiple r^2				0.34			

Table 11: Results from Forward Stepwise Multiple Regression Analysis

It will be evident from Table 11 above that Stepwise Multiple Regression accepted three steps. It will also be evident from the model presented above, that Mental Alertness explained the larger proportion (10.13%) of the variance in academic performance. Reading Comprehension and Conceptual Reasoning were responsible for smaller proportions of the variance. Reading Comprehension was responsible for 1.32% of the variance and Conceptual Reasoning for a non-significant, 0.56%. Thus, Conceptual Reasoning Test did not add much to the model. Although the r^2 is small (12.01%), Mental Alertness accounted for more than half of the variance explained.

It will be noted that the stepwise procedure did not include the Arithmetic Reasoning Test in the regression equation, indicating that this particular NIPR test did not make any significant contribution to predicting academic performance (defined as the average of the ten core courses in the MBA programme).

It was thus possible to conclude from the above regression analyses that Mental Alertness was the best single predictor of academic performance in the MBA programme, accounting for the

majority (10%) of the 12% proportion of variance in academic performance which the four NIPR tests were able to predict when used in combination.

6.4.1.2. Conclusions with Respect to Prediction of Academic Performance in the MBA Programme

The evidence presented up to this point in this chapter indicated a low predictive relationship between the four NIPR tests and academic performance, with Mental Alertness being the test which had most value as a selection tool. The results from the analyses presented above corroborated Barnard's earlier results (in J. M. Taylor, 1978). Barnard reported that Mental Alertness did not have a strong link with the courses measured in his study.

Barnard's criterion variables were the courses done in the first term. In the case of the WBS MBA, the complexities of the task of conducting selection were compounded by the length and variability of the courses taught as part of the curriculum. Lapse of time between predictor and criterion variables was also a problem. As Galpin (2001) has pointed out, interpretation of multiple regression becomes more complex in situations in which there is elapse of time, or time lag, between the administration of a test and the predictive criterion.

In the WBS MBA context, the core courses that formed the criterion variables in this analysis could be taken either over one year full-time, or over two years part-time. Even at this point in time (i.e. after completing the academic portion of the degree requirement), this did not mean that an MBA student had actually completed their studies. There was often a year to two more which elapsed before the research component was completed and the MBA qualification was completed. In many cases the research was not completed, and the students did not graduate.

Thus, it may well be that, as Galpin (2001) had pointed out, prediction studies done too soon or too late in one's studies tend to reveal negative or unusable results. It is normally ideal to run such studies using data of completion. In this study, this was achieved by considering a dichotomous completion/non-completion criterion variable, the point made is that the categories in this variable were not homogeneous, for the reason that certain students reached the stage of MBA degree completion after a few years and certain students did not reach the stage of degree completion at all. The results of this part of the analysis should thus be interpreted with caution.

Given a categorical (nominal) criterion variable, the second part of Research Question Two was answered through Logistic Regression, which is a form of regression designed to analyse data sets in which the dependent variable is dichotomous. Logistic regression predicts the logit of an event outcome from a set of predictors. In this study academic success was predicted from the four NIPR tests, namely, Mental Alertness, Reading Comprehension, Arithmetic Reasoning and Conceptual Reasoning as independent variables relative to completion/non-completion of the MBA degree as the dependent variable.

Results from this analysis are presented below. Before presenting these, the assumptions relating to the Logistic Regression procedure will be discussed.

6.4.2. Assumptions of Logistic Regression

Given the fact that the four NIPR tests were interval (continuous) variables, discriminant analysis could have been used in the analysis. Logistic regression was chosen as the statistical procedure of choice for the reason that it has several advantages over discriminant analysis (Galpin, 2001; Peng et al., 2002):

- It is more robust, as the independent variables do not have to be normally distributed, or have equal variance in each group,
- It does not assume a linear relationship between the Independent Variable and Dependent Variable,
- It can handle nonlinear effects,
- Explicit interaction and power terms can be added to the equation,
- The Dependent Variable does not need to be normally distributed,
- Normally distributed error terms are not assumed,
- It does not require that the Independent Variables be interval and
- It does not require that the Independent Variables be unbounded.

As Peng et al. (2002) suggest the procedure is a flexible one. "The only assumption that holds for logistic regression is that the binomial distribution is the assumed distribution for the conditional mean of the dichotomous outcome. This assumption implies that the same probability is maintained across the range of predictor values" (Peng et al., 2002, p. 11).

The binomial assumption is known to be robust as long as the sample is random, thus observations are independent from each other. Comments have already been made in the previous section concerning potential lack of homogeneity in the dependent variable in this study. The randomness of the sample was also not guaranteed. However compared to discriminant analysis, logistic regression was felt to offer the best available solution, and was thus the analytic procedure of choice.

Put another way, given the fact that the assumptions for multivariate regression were met in this particular data set (as reported earlier in this chapter), this implied that the less exacting

assumptions relating to logistic regression were also met. The procedure was thus applied to the case of calculating the probabilities relating to membership of the completion and non-completion categories of the MBA degree data as follows.

6.4.2.1. Results from Logistic Regression

A predictor logistic model was fitted to the data to test the hypothesis of whether or not the NIPR tests predicted academic success. Regrettably, none of the tests was successful in predicting the ratio of successful or non-successful completion of the MBA degree. According to the model below, the log of the odds of someone successfully completing their MBA was not related to the NIPR tests as no additional effects met the 0.05 significance level for entry into the model.

Response Profile		
Ordered Value	Success/Not	Total
		Frequency
1	0	312
2	1	429
Model Convergence Status		
Convergence criterion (GCONV=1E-8) satisfied.	=1008.693	
-2 Log L		

Residual Chi-Square Test		
Chi-Square	DF	Pr > Chi Sq
2.9893	4	0.5596

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > Chi Sq
Intercept	1	0.3185	0.0744	18.3184	<.0001

Profile Likelihood Confidence Interval for Parameters					
Parameter	Estimate	95% Confidence Limits			
Intercept	0.3185	0.1731	0.4649		

Table 12: Results from Logistic Regression Analysis

*** *No (additional) effects met the 0.05 significance level for entry into the model.*

The results from logistic regression in Table 12 above indicated that there were no significant differences between any of the four NIPR tests in predicting success versus non-success for the MBA degree (Chi Square=2.9893; p>0.5596). This can also be gauged from Table 13 below,

which reflects the means for each of the four NIPR tests relative to these two criterion groupings.

NIPR Test	Mean for the Successful	Mean for the Non-Successful
Mental Alertness	26.1	26.84
Reading Comprehension	15.87	16.06
Arithmetic Reasoning	17.55	18.3
Conceptual Reasoning	24.21	24.8

Table 13: Test Means of the Successful vs. Non-Successful Students with Reference to the NIPR Tests

6.4.2.2. Conclusions with Respect to Prediction of Academic Success in the MBA Programme

What the above analyses imply is that there were no significant differences in the test scores of the successful and unsuccessful groups. Surprisingly, the means of the scores of the unsuccessful candidates were slightly higher than those who have successfully completed their MBA studies. It was thus assumed that the inability of any of the four NIPR tests to predict academic success was based on the fact that the means were not significantly different for the two groups, implying that the distributions were not different. This in turn would imply that the odd ratios for the successful/not successful students would be similar or very close to 1 to indicate the even odd ratios (Galpin, 2001).

6.5. Results Relating to Research Question 3

- Do the four NIPR tests predict academic performance equally for the different racial and language groups enrolled in the WBS MBA programme?

- Do the four NIPR tests predict academic success equally for the different racial and language groups enrolled in the WBS MBA programme?

In this section, the analysis is extended to include race and language as additional variables. Some discussion has already been undertaken in the previous chapter concerning these two variables, and particularly the language variable which was operationalized as language of preference for purposes of communication, as opposed to home language spoken by the students.

Due to its complexity, the results relating to the language variable are reported first in the sections below, then in the results relating to the racial classification variable. The overarching research question was answered using the results yielded by the Stepwise Multiple Regression and Logistic Regression analyses discussed in the previous section, as well as by t-test and Analysis of Variance (ANOVA).

6.5.1. Results from Stepwise Multiple Regression with Language and Race as Classification Variables

6.5.1.1. Findings with Respect to Language Preference

In order to understand if the tests predicted equally for different language groups, multiple regression was conducted using language preference as a classification variable (See Appendix F1 for detailed results). The first stage of the analysis involved running separate exploratory analyses in which separate multiple regression analyses were conducted for students who

indicated that English was their preferred method of communication, and the students who indicated that they preferred communicating in languages other than English.

According to Appendix F1, for the 492 (66.4%) students that chose English as their preferred means of communication, r^2=11.33. Thus approximately 11% of the proportion of the variance in academic performance for this group was explained by the four NIPR tests (F=15.55; p<0.0001).

For the 249 (33.6%) students that indicated that they did not prefer English as a language of communication r^2=0.124 (F=8.7; p<0.0001). That means 12.4% of the variance in academic performance for this group was explained by the NIPR tests. Though the r^2 value was higher than the 11.3% variance explained for those who opted for English as a medium of communication, the proportion of common variance between predictors and criterion was still small.

Whether the difference of 1.15 in r^2 found between the two groups was significant and was tested post hoc, using ANOVA's and t tests. These analyses will be reported later in this section.

However, as the essential focus of Research Question 3 was to establish whether the four NIPR tests predicted academic performance better depending on the language preference of students, forward stepwise selection was utilized in order to establish which of the four NIPR tests was the best predictor in the case of each language and racial category of students.

	Summary of Forward Selection for English preference group							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	% of change	C(p)	F Value	Pr > F
1	MA	1	0.0874	0.0874	0.0874	13.1899	46.96	<.0001
2	RCT	2	0.0173	0.1048	0.0174	5.6806	9.46	0.0022
3	CRT	3	0.0068	0.1115	0.0067	3.9727	3.71	0.0547
4	ART	4	0.0018	0.1133	0.0018	5.0000	0.97	0.3245

Table 14: Results from Forward Stepwise Regression for English Speaking Students

It will be evident from table 14 above that Mental Alertness contributed the greatest proportion of the variance in academic performance in the English language preference group, with an r^2 of 8.7%. Though significant at $p<0.0001$, the percentage of common variance between this predictor and academic performance was not high. Adding the other NIPR tests into the equation added an additional 3% of variance predicted.

For the non-English language preference group, the results were as follows:

	Summary of Forward Selection for Non-English preference group							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	% of change	C(p)	F Value	Pr > F
1	MA	1	0.1020	0.1020	0.1020	5.36 40	28. 04	<. 0001
2	ART	2	0.0150	0.1170	0.015	3.18 21	4.1 8	0. 0420
3	RCT	3	0.0074	0.1243	0.0073	3.12 63	2.0 6	0. 1522

Table 15: Results from Forward Stepwise Regression for Non-English Speaking Students

This analysis indicated that the non-English group had a total r^2 of 12.4% with Mental Alertness explaining the majority (10.2%) of the proportion of variance in academic performance predicted. Arithmetic Reasoning and Reading Comprehension were responsible for an additional 2% of the variance predicted. Conceptual Reasoning was not accepted into the stepwise regression model.

It is evident from these results that Mental Alertness was the best individual predictor both for students who preferred communicating in English as well as those students who preferred communicating in other languages. The test explained 8% of the variance in academic performance for the English language preference group compared to 10% for the group that did not choose English as a medium of communication. All four tests were accepted into the model for the group that preferred English, whereas for the non-English group Conceptual Reasoning was not accepted into the model.

6.5.1.2. Findings with Respect to Racial Classification

A similar procedure was conducted with respect to analysis of the relationship between the four NIPR tests and racial classification. It will be recalled from the last chapter that the students were divided into three racial groupings, namely, White, Asian (Including both Indian and Chinese students) and Black (including both African and so-called Coloured) students. The first stage of the analysis involved running separate exploratory analyses in which separate multiple regression analyses were conducted for each of these samples of students.

The results pertaining to the relationship between the four NIPR tests and the academic performance of the Asian group of MBA students are presented in the table following and the detailed ones are presented in Appendix F2:

	Summary of Forward Selection for the Asian Group							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	% of change	C(p)	F Value	Pr > F
1	MA	1	0.0343	0.0343	0.0343	0.8880	3.84	0.0528
2	ART	2	0.0098	0.0441	0.0098	1.8037	1.10	0.2974
3	RCT	3	0.0071	0.0512	0.0071	3.0147	0.80	0.3742

Table 16: Results from Forward Stepwise Regression for Asian Students

From Table 16, it will be noted that for the Asian group, Mental Alertness was the best predictor of academic performance, followed by Arithmetic Reasoning then Reading Comprehension. Conceptual Reasoning was not accepted by the model. It is however important to comment about the magnitude of the proportions. The 3.4% common variance predicted by Mental

Alertness is very small, in terms of the proportion that is explained by the best predictor in a regression equation. A total of 5% common variance explained by all the NIPR tests is also very low.

Nor were the results pertaining to the relationship between the four NIPR tests and the academic performance of the Black group of MBA students much better, as will be evident from the results presented in the table following.

		Summary of Forward Selection for Black Group						
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	% of change	C(p)	F Value	Pr > F
1	MA	1	0.0512	0.0512	0.0512	1.5914	9.72	0.0021
2	RCT	2	0.0074	0.0586	0.0074	2.1948	1.40	0.2378
3	ART	3	0.0063	0.0649	0.0063	3.0115	1.19	0.2768

Table 17: Results from Forward Stepwise Regression for Black Students

The trends from this analysis were similar to the trends found for the Asian group where Mental Alertness was the best single predictor of academic performance for the Black group of students, explaining 5.1% of the common variance between the two variables. Overall, the combination of the NIPR tests increased the variance predicted to 6.5%. As in the analysis for the Asian students, Conceptual Reasoning was not accepted into the model.

Overall, these results indicated that the NIPR tests had a very low predictive relationship with the academic performance for both these groups of students. This was also the case for the White students as will be evident from the results presented in the next table.

	Summary of Forward Selection for White Group							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	% of change	C(p)	F Value	Pr > F
1	RCT	1	0.0593	0.0593	0.0593	12.1729	28.2	<.0001
2	CRT	2	0.0229	0.0822	0.0229	3.0458	11.13	0.0009
3	ART	3	0.0031	0.0853	0.0031	3.559	1.49	0.2231
4	MA	4	0.0012	0.0864	0.0011	5	0.56	0.4551

Table 18: Results from Forward Stepwise Regression for White Students

It will be noted from table 18 above that all the four NIPR tests were accepted into the regression model for White students. Reading Comprehension was the best single predictor, accounting for 6% of the variance in academic performance. Conceptual Reasoning accounted for a further 2% of variance in the model. In contrast to the Black and Asian students, Mental Alertness added very little in terms of additional variance to the regression equation and was not the best predictor as had been found with both the Asian and Black students. Instead the test was the lowest contributor to the variance in academic performance.

Thus for White students, Reading Comprehension was the best predictor followed by Conceptual Reasoning then Arithmetic Reasoning and finally Mental Alertness which contributed only 0.0012% additional variance to the total 8.16% variance predicted by the NIPR test battery as a whole.

6.5.2. Discussion

Given the literature in the area, there were some surprising findings from the analysis. One surprising finding relative to the literature (e.g., J. M. Taylor, 1980, 1981) was that Mental Alertness was not the best predictor for White students. However, it was for the Asian and Black students. Another was that Conceptual Reasoning was not accepted into the models for either the Asian and Black students, but it was the second best predictor of academic performance for White students.

The Conceptual Reasoning results came as a surprise to the researcher considering literature debates around the usability and effectiveness of non-verbal tools for Black students (Boeyens, 1990). The Conceptual Reasoning test is classed as such, however it did not emerge as an important test in determining academic performance for Black students in this study.

However, conflicting results were presented by researchers in Erwee (1981)'s study about the difficulties experienced by Black students and employees in certain situation that require changing their strategy because a culturally conditioned bias dictates that the perceptual mode is practical and useful in an African society. Thus, a society that favours concrete perceptual strategy than more abstract-conceptual approaches. However, "in order to advance to higher level positions Black candidates will be forced to shift their focus" (p. 30). "An increasing number of Blacks in South Africa are becoming available for jobs in industry and commerce which involve more than mere routine" (p. 31). Hence the increasing demands for MBA studies.

Non-routinized jobs demand moderate to high levels of abstract-conceptual functioning. Even though Black student's upbringing does not give them an opportunity to play around this trait, J. M. Taylor (1977) acknowledged that with practice and increased familiarity students were able

to shift from a primarily concrete perceptual approach to a more abstract-conceptual approach in the solution of test items.

Another unforeseen outcome was that of Mental Alertness for White students. Many studies that have been administered on the reliability and validity of the NIPR tests have found Mental Alertness to be the better (amongst the four) predictor of academic performance. J. M. Taylor's (1978) study confirmed this finding with White English and Afrikaans speaking students. The results for Black and Asian students confirmed this theory, however the contradiction was observed with White students.

The findings with respect to the language variable are difficult to interpret, for the reason that the variable examined was English language preference for purposes of communication. The Reading Comprehension Test (which is effectively a test of competence in written language) was found to be the best predictor of academic performance for White student and was the second best predictor for that particular group. However for Black students it was the third best predictor and explained a very small proportion of the variance in academic performance.

Given these apparent differences in the ability of the four NIPR tests to predict academic performance, additional analyses were conducted to establish whether the four NIPR tests predicted academic success, as defined by the successful and failure to complete MBA studies offered by the Wits Business School. Following the logic already presented earlier in this chapter, Logistic Regression was used as the relevant statistical procedure for this aspect of the analysis. Race and language then served as classification variables.

6.5.3. Results from Logistic Regression with Language and Race as Classification Variables

6.5.3.1. Findings with Respect to Language Preference

None of the tests were accepted into the model as predictors of academic success (See Appendix G). Although the test scores for the group that did not prefer English were higher than those of the candidates who preferred English, the odd ratios of being in one or other of these groups was not very high. The odds of successfully completing MBA studies given English as a first language is then calculated as p/ (1-p) that is using data in table 19:

[288/492] / [1-288/492] = [288/492] / [204/492]

= 288/204

so that the odds are 1.4118 to 1 (Galpin, 2001).

Furthermore, the odds of successfully completing the MBA degree given that English is not the first language preferred for communication are 1.0814. Therefore, students speaking English as a first language have odds of 1.0814 to 1 of completing their MBA studies compared to those who opted for other languages.

Variable	Successful	Failed to Succeed	Total
English	288	204	492
Non-English	141	108	249
Total	429	312	741

Table 19: Odd Ratios Table for Language Groups

"Since one would seldom get an odds ratio of exactly 1, one needs a method to test whether an odds ratio is significantly different from 1" (Galpin, 2001, p. 213). This is done by calculating a standard error.

The standard error for the log odds ratio above is presented by the following formula:

$=\sqrt{[(1/288) + (1/204) + (1/141) + (1/108)]}$

$=\sqrt{0.0247}$

=0.157, thus standard error is significantly different to 1

What these results indicated was that the overall logistic model was not successful in predicting these overall group differences, and also that the variance in the four NIPR tests did not explain the odd ratios of being successful for students who preferred communicating in English versus those who preferred communicating in other languages. Whether there were significant differences in the means and distributions of the individual NIPR tests was then also tested post hoc using t-tests and ANOVAS, as will be reported in the sections following.

6.5.3.2. Findings with Respect to Racial Classification

Similar results using logistic regression to test for differences in racial classification were found to the results for language reported above. None of the NIPR tests was able to predict academic success for the three racial groups comprising the racial classification variable (White, Asian and Black) (Refer to Appendix I).

It is unfortunate because what all this means is that none of the tests could predict academic success for the WBS MBA programme.

Variable	Successfully completed	Failed to complete	Total
White	284	165	449
Black	86	96	182
Asian	59	51	110
Total	429	312	741

Table 20: Odd Ratios Table for Racial Groups

The results of the analysis indicated that the overall odds of successfully completing the MBA qualification were 1.375 to 1 (429/741 to 312/741). There are more chances of completing the degree. But with values very close to 1, the odds are more or less even. This means the odds of succeeding are more or less equal or the same as the odds of failing to successfully complete one's MBA studies.

The odds of Whites students (compared to Black students) successfully completing their MBA studies are 1.3386 to 1. Thus White students are more likely than Black students to successfully complete the WBS MBA.

The calculation above is based on the following formula:

284/429 to 86/182=1.3386

The standard error of the log odds ratio above is presented by the following formula:

$=\sqrt{[(1/284) + (1/165) + (1/86) + (1/96)]}$

$=\sqrt{0.0316}$

=0.177, thus significantly different to 1.

Given the indications of differences within the odds for different racial groups relative to the overall criterion of successful completion of the MBA degree, the next phase in the analysis involved comparing means of the language and racial groups in the study. This analysis was undertaken using either t-test or ANOVA depending on whether the demographic variable had two or more levels.

6.5.4. Results from Post Hoc Testing of Differences within the Language Variable

	Variable	N	Mean	Std Dev	T-value	P value
Mental Alertness	**Language 1**	492	26.9797	6.4432	3.27	0.0011
	Language 2	249	25.3333	6.5591		
Reading Comprehension	**Language 1**	492	16.2175	4.1865	2.5	0.0128
	Language 2	249	15.4378	3.6584		
Arithmetic Reasoning	**Language 1**	492	18.4654	6.6526	3.4	0.0007
	Language 2	249	16.7189	6.4871		
Conceptual Reasoning	**Language 1**	492	24.8841	5.8617	2.78	0.0056
	Language 2	249	23.6225	5.7797		
Academic Performance	**Language 1**	492	66.628	4.0866	5	<.0001
	Language 2	249	65.0197	4.2287		

Table 21: Results from t-tests

It will be noted from Table 21 that significant differences were found in the scores of all the four NIPR tests for the two different language groups, with those students who preferred using English for communication scoring significantly higher than those students who preferred using other languages for purposes of communication. All results were significantly well beyond the critical level of $p<0.05$. Given this, Bonferroni statistics were not calculated but effect sizes were estimated as follows:

Variable	Means	Std Deviations	Cohen's D	Effect size r
Mental Alertness	29.98 25.33	6.44 6.56	0.25	0.13
Reading Comprehension	16.22 15.44	4.19 3.66	0.20	0.10
Arithmetic Reasoning	18.47 16.72	6.65 6.49	0.27	0.13
Conceptual Reasoning	24.88 23.62	5.86 5.78	0.21	0.11
Average Core Courses	66.63 65.02	4.09 4.23	0.39	0.19

Table 22: Details of Effect Sizes for the Two Language Groups

Following Cohen (1988), for Cohen's d an effect size of 0.2 to 0.3 might be a "small" effect, around 0.5 a "medium" effect and 0.8 to infinity, a "large" effect. In terms of these guidelines the effect size for the differences referred to in tables 21 and 22 above would be:

Mental Alertness: small

Reading Comprehension: small

Arithmetic Reasoning: small

Conceptual Reasoning: small

Overall Academic Performance: small

As Cohen comments (1988, p. 25): "the terms 'small,' 'medium,' and 'large' are relative not only to each other but to the area of behavioural science or even more particularly to the specific content and research method being employed in any given investigation....In the face of this

relativity, there is a certain risk inherent in offering conventional operational definitions for these terms for use in power analysis in as diverse a field of inquiry as behavioural science".

Overall the analyses presented in this section provide indications that significant differences existed between the performances of the two different language groups as measured through the four NIPR tests, as will be evident in Appendix H. The effect sizes of these differences were, however small.

This result should be interpreted with caution as the researcher had already suggested in Chapter Four that the language variable may have been contaminated due to the fact that the non-English variable included students speaking a wide variety of languages as their mother tongues. Therefore the interpretation of this variable may be affected by this potential contamination.

Nevertheless, the results from these analyses indicated that language may act as a barrier to learning and performance as indicated in the literature by Broom and Doctor (1994) MacFarlane (2006).

6.5.5. Results from Post Hoc Testing of Differences within the Racial Classification Variable

For this analysis ANOVA was used for the reason that the independent variable had three levels (See Appendix I for a detailed presentation of ANOVA results).

Level of Race	N	MA		RCT		ART		CRT		Average Core Courses	
		Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Asian	110	26.982	5.6406	15.564	3.8541	18.564	6.58818	24.2818	5.71355	66.3909	3.48848
Black	182	22.769	6.1205	14.665	4.0114	15.549	6.56167	21.6538	5.63983	63.5346	3.86232
White	449	27.773	6.3335	16.575	3.9533	18.655	6.48444	25.6414	5.598	67.0481	4.07114

Table 23: NIPR Tests and Academic Performance Distributions by Race

According to table 23 above, there was a difference between the scores of the three racial groups in their performance on the four NIPR tests, with Black students' scores (e.g. on the Mental Alertness test) lower than those of the White and Asian students. On the side of the dependent variable, there were similar differences. Here the difference between the scores on the White and Asian students were also not large (a difference of 0.66 between the mean of 66.39 obtained by Asian students and the mean of 67.05 obtained by White students). However there was a larger difference observed between mean score of Black students (63.53) and students in the other two racial categories. What this meant was that there was an observed difference of 3.51 between the mean academic performance of Black and White students and an observed difference of 2.86 between the mean academic performance of Black and Asian students. The significance levels applying to these differences are presented in Table 24.

Source of variation	Df	Sum of Squares	Mean of Squares	F-ratio	P-value
Mental Alertness Test	2	3282.142	1641.1	42.92	<0.0001
Error	738	28219.1	38.237		
Corrected Total	740	31501.24			
Reading Comprehension Test	2	492.1703	246.09	15.75	<0.0001
Error	738	11533.36	15.628		
Corrected Total	740	12025			
Arithmetic Reasoning Test	2	1309.467	654.73	15.41	<0.0001
Error	738	31361.6	42.495		
Corrected Total	740	32671.07			
Conceptual Reasoning Test	2	2063.35	1031.7	32.6	<0.0001
Error	738	23354.73	31.646		
Corrected Total	740	25418.08			
Academic Performance	2	1610.582	805.29	51.9	<0.0001
Error	738	11451.78	15.517		
Corrected Total	740	13062.37			

Table 24: Results from ANOVA Analysis

Given that the significant levels for all NIPR tests were well beyond both the 0.05 and 0.01 levels, the conclusion was that there was a significant difference both in performance on the four NIPR tests as well as a significant difference in the academic performance for the different races in the study.

The results showing significant differences existed between the academic performances of the different racial groups in the study (F (2.738) =51.9; p= (0.0001)) were then explored further. Post hoc Tukey's testing showed that White students (mean=67.05%) differed from both Black (mean=63.5%) and Asian students (mean=66.39%). White students scored significantly higher than the Black student group. Asian students also scored significantly higher than the Black student group.

In order to determine effect size using ANOVA, an index which Cohen (1988) has referred to as η^2 (Eta squared) rather than Cohen's d, was used. According to Cohen (1988) Eta squared values of:

- 0.01: Small
- 0.059: Medium
- 0.138: Large

What this means is that if one ends up with $\eta^2 = 0.45$ the effect size is very large. It also means that 45% of the change in the DV can be accounted for by the IV.

The formula used for calculating eta squared is:

$$\eta^2 = \frac{\text{Treatment Sum of Squares}}{\text{Total Sum of Squares}}$$

In this study, effect size according to Mental Alertness scores:

$$= \frac{3282.142}{31501.24}$$

=0.1042

This means that 10.4% of the variance in academic performance of the difference race groups in the study can be accounted for by the Mental Alertness scores. This, according to Cohen's guide is large.

A detailed representation of the eta squared for all NIPR tests is presented in table 25 below. All, except Mental Alertness and the Average of the Core Courses, producing medium effect sizes. The effect sizes for both Mental Alertness and the Average of the Core Courses (the academic performance variable) were large.

Variable	Sum of Squares	Corrected Total	Eta Squared
Mental Alertness	3282.142	31501.24	0.1042
Reading Comprehension	492.1703	12025.53	0.041
Arithmetic Reasoning	1309.467	32671.07	0.0401
Conceptual Reasoning	2063.35	25418.08	0.081
Average Core Courses	1610.582	13062.37	0.1233

Table 25: Details of Effect Sizes for the Three Racial Groups

6.6. Conclusion

This chapter has presented findings of various analyses. Tests of correlations between the four NIPR tests forming the independent variables and academic performance as the dependent variable revealed some significant relationships, but overall the levels of relationship were not firm. The multiple correlation between the four NIPR independent variables and the two criteria of academic performance and academic success also did not indicate a strong fit. The r^2 for the tests combined explained around 12% of the proportion of the variance in academic performance. This meant that 88% of the variance in academic success as defined by the average core courses was explained by third variables unknown to the researcher.

Stepwise selection multiple regression revealed that overall Mental Alertness was the best predictor of academic performance explaining at least 80% of the total variance predicted. This test was followed by Reading Comprehension and Conceptual Reasoning. However the difference added by these two tests was very small. Arithmetic Reasoning was the only test that was not accepted into the model. Hence, it did not serve as an important predictor of overall academic performance for the whole sample. Logistic regression showed that none of the variables were significant in predicting whether students would successfully complete their MBA studies or fail to complete their MBA studies. None of the tests were accepted into the model.

The last question in this study focused on whether tests predicted academic performance and academic success differently for different racial and language groups. The results revealed a number of differences between language groups. Some significant differences (3.5135) were also observed in the results of Black and White students (*M*= 63.5% and 67.05% respectively) and in the results of Black and Asian students (*M*= 63.5% and 66.39% respectively) but not between the results of the Asian and White students (*M*= 66.39% and 67.15% respectively).

Regression analyses were also conducted for the different language and racial groups. There were a number of differences between these groups. The most intriguing one was for White students where Mental Alertness was not the strongest predictor as was found in the analyses conducted for Black and Asian students. For White students Reading Comprehension was the best predictor. Conceptual Reasoning on the other hand emerged as a non-important predictor for Black and Asian students but was the second important predictor of academic performance for White students.

With regards to the prediction of academic success and the differences in scores of the different racial groups in the study the results revealed that there was no difference in how the tests predicted academic success for the different groups. This finding stems from the fact that none of the four NIPR the tests were accepted into the logistic regression model in predicting academic success.

However post hoc testing of the differences between the scores obtained by the different racial groupings in the data set provided indications that significant differences existed between the performances of the three racial groups measured through the four NIPR tests, as will be evident in Appendix I. The effect sizes of these differences were estimated as medium to large confirming the significant differences.

7. Conclusions, Limitations and Recommendations for Future Studies

7.1. Introduction

This research focused on the ability of the four NIPR tests in predicting academic performance and academic success in the Wits Business School's MBA programme. Focus was also placed on the tests' ability to predict equally for different racial and language groups in the study.

The research addressed the following research questions:

- Are there significant relationships between the four NIPR tests and academic performance in the WBS MBA curriculum?

- Do the first order correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?

- Do the multiple correlations indicate that the four NIPR tests predict academic performance in the WBS MBA curriculum?

- Do the four NIPR tests predict academic success in the WBS MBA programme?

- Do the four NIPR tests predict academic performance equally for the different racial and language groups enrolled in the WBS MBA programme?

- Do the four NIPR tests predict academic success equally for the different racial and language groups enrolled in the WBS MBA programme?

7.2. Summary of Key Findings in the Study

In this research, the scores on four NIPR tests (the Mental Alertness, Arithmetic Reasoning, Conceptual Reasoning and Reading Comprehension tests) formed the independent variables and course results on the MBA programme the dependent variables. The analysis involved examining the relationship between the four NIPR tests and academic performance in the ten core courses of the MBA as one dependent variable and academic success in completing the MBA degree as a second dependent variable. The language background and racial classification of the students were also examined in relation to academic performance and academic success.

Using intercorrelation it was found that significant and moderately firm first order relationships existed between the four NIPR tests forming the independent variables in the study. In terms of inter-relationships between the core courses forming part of the academic performance criterion variable, it was found that a firm relationship existed between Data and Decision Making together with Finance/Financial Management where r=0.524. Operations Technology was the only course that had significant yet firm relationships with most of the other MBA core courses.

When the multivariate relationship between the independent and dependent variables was examined, the results of stepwise regression indicated that Mental Alertness was the best predictor of academic performance overall but that it explained only 10% of the variance in academic performance. Adding the additional predictor variables into the regression equation added only an additional 2% to the amount of variance which could be predicted through use of

the four NIPR tests in combination. Logistic regression showed that none of the variables were significant in predicting whether students would successfully complete their MBA studies or fail to. None of the four NIPR tests were able to meet the significance level necessary to be entered into the model.

Separate regression analyses were also conducted for the different language and racial groups in the study which indicated a number of differences between these groups. The most important findings from this part of the analysis was that for White students, Mental Alertness was not the strongest predictor as was found in the analyses conducted for Black and Asian students. For White students Reading Comprehension was the best predictor and Conceptual Reasoning the second important predictor of academic performance.

Conceptual Reasoning in contrast, was not a significant predictor in the stepwise regression model for Black and Asian students. Overall the proportion of variance explained by Mental Alertness for the Black and Asian students was very small standing at 5% and 3.4%, respectively.

Although Mental Alertness was not the best predictor of the academic performance for White students, it was the best predictor (8.74%) of the variance in academic performance for the group that selected English as the preferred language of communication. Reading Comprehension was the second most important predictor for this group contributing an additional 1.74% of common variance predicted, followed by Conceptual Reasoning adding 0.67% to the model and then Arithmetic Reasoning adding a mere 0.18% of variance to the regression model.

Given the above indications of differences within the overall sample, additional post hoc analyses were conducted which indicated a number of differences in academic performance which could be attributed to language preference. It was evident that the group that preferred English had higher odd ratios than the group that did not prefer English as a language of communication. However the effect sizes of these differences were small. It should be noted that the language variable was defined in terms of differences in language preference as opposed to differences in the home languages spoken by the MBA student sample. As such these differences might not be a true reflection of the sample's language characteristics.

With regards to differences between the academic performances of the different racial groups it was evident from the analyses that Black students were outperformed by White and Asian students. The difference between Asian and White mean scores was very small compared to the difference between mean scores of Black and White students. The effect sizes of these differences were medium to large. In terms of differences in the academic success of different racial groups, White students had better odds of successfully completing their MBA studies than Black students. In addition, Asian students also had better odds of succeeding than Black students.

The overall conclusion from this study was that while the abilities tapped by all four NIPR tests contributed towards determining academic performance, their overall contributions were not high. They could only account for 12% of the variation in academic performance of the Wits MBA student sample. In contrast, extraneous variables explained 88% of the variance in academic performance. This indicated that the four NIPR tests contributed information which was only marginally useful to the selection decision.

7.3. Strengths of the Study

Sample Size

The large sample size was the key strength of the study. Although this study did not adopt a probability random selection approach, the sample size of 741 was large enough to infer some of the findings to the MBA class at large.

7.4. Limitations of the Study

Some limitations have already been addressed in the report this section however gives a detailed analysis of these limitations.

The major limitation in this research related to sampling and stemmed from the fact that there were missing data in the archival records from which the data set for this study was constructed. There were over 2000 records available but a sample of 741 student records was constructed for analysis. Those records that were not complete due to students being either exempted from particular course requirements or dropping out of the programme were not able to be used. This meant that the sample of data ultimately analysed was an attenuated one and comprised only the records of the students which were complete. This limitation may have affected the results reported from this research.

Another limitation applied to the language variable which was developed as categorical data comprising either students who expressed a preference for communicating in English and those students who preferred communicating in other languages. Similar limitations applied to the racial categorisation variable in which the results of Indian and Chinese students were elided

into a single Asian student category and the results of African and so-called Coloured students were elided into a single Black student category.

These limitations are likely to have had the following influence on the validity of the results reported above.

External Validity

The method of selecting the sample for this study was not random and it involved some judgement of the available data to select the participants who fit the requirements of the current study. Therefore, generalisability of the results from this study to the general MBA students is limited. It cannot be confidently concluded that this study is completely representative of the WBS MBA class, as well it cannot be inferred that the findings from this study can be generalised to the population of MBA students in South Africa. However the findings can be used as a guide for decision making.

Maturation

The period between the administration of the tests and the completion of the MBA course is far apart which presents a substantial duration for the occurrence of processes that may have resulted in the changes in the subjects. It takes a minimum of three years to complete a part time MBA (WBS Information, 2007). The time delay between the first phase and the final one may have introduced other influences into the study that may have confounded the phenomenon observed and the results obtained (Potter, 1998). Unfortunately the researcher had no control over these potential third variables.

Confounding Variables

There is no guarantee that no one has manipulated the data since it has been captured in the University's archives. Due to the confidentiality of the study, students will not be given an opportunity to confirm the accuracy of their academic course scores. The researcher had to accept the information at face value.

The university introduced the Oracle system in 2006 and this was a concern for Dr Swemmer (University Registrar) when the researcher requested to undertake this study. He mentioned that with the move from the old system to the new one there is no guarantee that there will not be an interference with the student records (Dr. D. Swemmer, personal communication, February 17, 2007).

History

Similar to the confounding variables history looks at the effect of environmental events over time. History acknowledges the fact that those other environmental events, which are not the variables of interest in this study could have an impact, positive or negative, on the subject and can be erroneously calculated as a variable of study (Dockrat, 2005).

Research Approach

The sole use of the Quantitative approach can be regarded as a limiting factor in terms of the information received. In order to expand the study further and get more quality information, the study could have considered obtaining responses from a small sample of students to establish their views on the factors that predicted academic success for them.

Instruments

The use of tests for selection purposes is also challenging in that tests show what an individual can do at a given time but they fail to explain why a person is behaving the way they do.

Violation of the normal distribution requirement

It is expected that when dealing with bigger samples one would have to deal with issues of non-normal distribution of data. Though this was a minor obstacle in understanding the data and coming up with the relevant statistics, it should be acknowledged that they could have had an effect on the coefficients and overall statistics observed. The researcher did however control over these effects by running outlier analysis.

Informed consent

The study gets into students' confidential records and the consent of the involved candidates was not obtained. Instead the researcher obtained authorisation to conduct the research from the Head of Student Affairs, the then Head of the WBS and the then Head of Academic Programmes at the WBS.

Secondary data

Using data that already exists can cause problems in terms of restricting one's options. As a researcher one has to work with what one has, there is no room for creativity. One would have liked to have explored other variables, especially the non-cognitive ones. However such information was not collected hence the limitation for the researcher.

7.5. Suggestions for Future Research

It is evident from the statistical analysis that the four NIPR tests are responsible for a small proportion of the variance in academic performance. It is however important that future studies explore other variables that might have an influence on the prediction of academic performance and academic success in the WBS MBA programme. Such variables can be non-cognitive or a combination of cognitive and non-cognitive variables like in Payne's study (learning styles, motivation, anxiety, and so forth). Randsell (2001) confirmed that cognitive and non-cognitive variables are equally responsible for the variance in academic performance and academic success.

Another suggestion for future studies is the combination of quantitative and qualitative analysis like in Mokone's (2008) study. After running quantitative analysis, it would be ideal to confirm and explore results through additional qualitative analysis. This way an in-depth and holistic view of the variables would have been achieved.

7.6. Conclusion

The overall conclusion from this study is that, the four NIPR tests used as a selection test battery for the Wits MBA programme did not have a firm relationship with either subsequent academic performance or with academic success in terms of ultimately graduating in the degree. The continued use of this particular test battery must therefore be questioned. If the WBS considers reintroducing the NIPR tests (as these were discontinued at the end of 2009), there would be some evidence supporting a decision to keep the Mental Alertness in the selection test battery as the test came out as the best predictor of academic performance overall and for Black and Asian students.

Essentially, the argument would be that although the proportion of variance predicted by the Mental Alertness test was small for Black, Asian and White students, the evidence from this study indicated that it had many significant first order correlations with a number of the criterion course variables in the MBA programme. Reading comprehension was also a significant predictor especially for White students in the MBA programme.

Seeing that Conceptual Reasoning was not accepted into either the Black and Asian stepwise regression models and considering the legislative pressures that institutions of higher education face with regards to reformation and transformation of their racial representation, it would be recommended that the Conceptual Reasoning test be eliminated from the WBS selection test battery. Although the Conceptual Reasoning test might be eliminated, Mental Alertness would still address the inductive reasoning trait. The reason being a high correlation existed between Mental Alertness and Conceptual Reasoning Test (r=0.65).

Arithmetic Reasoning test was not accepted by the overall stepwise regression model, however using classification variables the test was entered and accepted into the models for the three racial groups and two language groups. Nonetheless, it should be noted that the variance in academic performance explained by this test was very low. It is due to these findings that the researcher suggests that the use of the Arithmetic Reasoning Test be reviewed further with a different sample focusing on other demographic variables.

Overall the use of the four NIPR tests would appear to have a limited usefulness for selection purposes for this particular degree programme. Use of the tests in combination as a selection test battery may need to be re-evaluated with all stakeholders involved in the selection process.

References List

Allison, P. D. (2006). *Logistic regression using SAS: Theory and application*. USA: SAS institute and Wiley.

Alkin, M. C. (1969). Evaluation theory development. *Evaluation Comment, 2, 1*, 2-7.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (*7th ed.*). Upper Saddle River NJ: Prentice Hall.

Ardila, A. (2001). Predictors of university academic performance in Colombia. *International Journal of Education Research, 35,* 411-417.

Beukes, D. P. M. (1969). *High level battery: Test administrators manual* (*3rd ed.*). Johannesburg: NIPR.

Bezuidenhout, P. U. (1997). *The role of MBA programmes within management education.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Biesheuvel, S., & Liddicoat, R., (1958). The effects of cultural factors on intelligence test performance. *Journal of the National Institute for Personnel Research, 8,* 3-14.

Boeyens, J. C. A. (1990). *The manual for the conceptual reasoning test.* Pretoria: National Institute for Personnel Research.

Bordens, K. S., & Abbott, B. B. (2008). *Research design and methods: A process approach* (*8th ed.*) New York: McGraw-Hill.

Breytenbach, J. W. (2008). *Finding statistical models using psychometric tests, matric results and biographical data to predict academic success at a South African university*. Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Bricker, R. I. (1990). *Management development requirement in a changing South Africa.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Broom, Y. M., & Doctor, E. A. (1994). Developmental dyslexia: Contrasting patterns of performance on a diagnostic psycholinguistic assessment. *South African Journal of Psychology. 24, 4*, 219-225.

Buntin, I. (1994). *A legacy of inequality: Higher education in South Africa*. Cape Town: UCT Press.

Carmichael, T. (2003). *The perceived return on investment of an MBA*. Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Code of Practice for Psychological Assessment for the Workplace in South Africa (1998). July 1998. www.pai.org.za/codeof.html 07/09/2000.

Cockcroft, K., & Laher, S. (in press). Current Trends in Psychological Assessment in South Africa: Challenges and Opportunities. In R. Basson, P. Fridjhon, S. Porter & C.S. Potter (Eds.), *Monitoring and evaluation in developing world contexts.* A Virtual Conference Publication of the

Wits Programme Evaluation Group. http://wpeg.wits.ac.za with an introduction by David M. Fetterman. Johannesburg: Malagasy Press. Publication April 2012.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd *ed.*). New York: Lawrence Erlbaum Associates.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioural sciences.* New Jersey: Lawrence Erlbaum Associates, Publishers.

Cook, J. (year not specified). *An informal study on the reliability and validity of the NIPR tests as utilized at the Wits Business School.* Johannesburg: University of the Witwatersrand.

Cook, J., & Owens, J. (1991). *Strategies for success in graduate studies.* Unpublished report. University of the Witwatersrand, Johannesburg.

Cooper, J. D. (1986). *Improving reading comprehension.* Boston: Houghton Mifflin Co.

Cronbach, L. J. (1960). *Essentials of psychological testing* (2nd *ed.*). New York: Harper and Bro.

Culverwell, D. (1989). *Learning strategies and academic success of undergraduate arts students.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Dale, G. J. (2010). *Predicting learner performance in clothing industry.* Unpublished PhD thesis. Cape Peninsula University of Technology, Cape Town.

De Beer, M. (2005). *Learning potential computerised adaptive test: User's manual.* Mucklenuek: Production Printers.

Dockrat, S. Y. (2005). *Selection for remedial intervention: The validation of the academic proficiency assessment battery as a predictor of academic achievement.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Doody, D. J. (1997). *An evaluation of the Wits MBA programme in satisfying graduates educational requirements and expectations.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Elkin, G. R., & Sneyd, C. (1989). The relationship between performance on the graduate management admission tests and academic results for graduate management students. *New Zealand Journal of Psychology, 18*, 82-83.

Erwee, R. (1981). Cognitive functioning, achievement motivation and vocational preferences of Black university students. *Psychologia Africana, 20*, 29-51.

Eysenck, M. W. (1993). *Principles of cognitive psychology.* Howe, UK: Lawrence Erlbaum Associates Publishers.

Fish, L. A., & Wilson, F. S. (2007). Predicting performance of one year MBA students. *College Student Journal, 41, 3*, 507-514.

Foxcroft, C. D., & Roodt, G. (2001). *An introduction to psychological assessment in the South African context.* Cape Town: Oxford University Press.

Foxcroft, C. D. (in press). Some issues in assessment in a developing world context: An African perspective. In R. Basson, P. Fridjhon, S. Porter & C.S. Potter (Eds.), *Monitoring and evaluation in developing world contexts.* A Virtual Conference Publication of the Wits Programme Evaluation Group http://wpeg.wits.ac.za with an introduction by David M. Fetterman. Johannesburg: Malagasy Press. Publication April 2012.

Fraser, W. J., & Killen, R. (2003). Factors influencing academic success or failure of first year and senior students: Do educators and lectures perceive things differently? *South African Journal of Education, 23, 4,* 254-260.

Frisby, C. L., & Braden, J. P. (1992). Feuerstein dynamic assessment approach: A semantic, logical and empirical critique. *Journal of Special Education, 26, 3*, 281-301.

Galpin, J. (2001). *Statistical research design and analysis handbook*. Department of Statistics and Actuarial Science, University of the Witwatersrand, South Africa.

Gaydon, V. P. (1988). *Predictors of performance of disadvantaged adolescents on the Soweto/Alexandra gifted child programme*. Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Gewer, A. (1998). *Uncovering potential: Dynamic assessment of non-verbal reasoning ability in educationally disadvantaged children.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Gregory, R. J. (2007). *Psychological testing: History, principles and applications* (5^{th} *ed.*). Boston: Pearson.

Grellet, F. (1981). *Developing reading skills. A practical guide to reading comprehension exercises.* London: Cambridge University Press.

Halstead, M. E. (1985). *The development and validation of three tests of arithmetic ability.* Pretoria: Human Sciences Research Council.

Hoffenberg, S. R. (1988). *Effectiveness of the learning potential assessment device with high achieving adolescents from a disadvantaged community*. Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Holburn, P. T. (1991a). *Differential item functioning in the mental alertness test*. Unpublished master's dissertation. University of South Africa, Pretoria.

Holburn, P. T. (1991b). *Selection decisions: The quest for fairness*. An NTB HSRC Report. Pretoria: Human Sciences Research Council.

Holburn, P. T. (1992). *Test bias in the intermediate mental alertness, mechanical comprehension, blox and high level figure classification tests*. An NTB HSRC Report. Pretoria: Human Sciences Research Council.

Howell, D. C. (2008). *Fundamental Statistics for the behavioral sciences* (6^{th} *ed.*). Belmont: Thompson Publishers.

http://www.wits.ac.za/depts/psychology.

http://www.uct.ac.za/depts/graduateschoolofmanagement.

Huysamen, G. K. (1980). *Psychological testing in South Africa*. Pretoria: Academica.

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

Huysamen, G. K. (1994). *Methodology for the social and behavioural sciences.* Halfway House: Southern Books Publishers.

Israel, N. (2006). *Raven's advanced progressive matrices within a South African Context.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Jensen, A. R. (1973). *Educability and group differences*. London: Methuen and Company Ltd.

Kagan, J. (2003). *MBA South Africa: A strategic guide to selecting your MBA programme and school.* Cape Town: Zebra Press.

Kachigan, S. K. (1991). *Multivariate statistical analysis: A conceptual introduction* (2^{nd} *ed.*). New York: Radius Press.

Kaufman, W. (2003). *An evaluation of female engineering student's academic performance in first year engineering graphics.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Kendall, I. M. (1981). *Psychological testing in the context of personnel selection.* Johannesburg: National Institute for Personnel Research.

Killen, R., Marais, A., & Loedoff, P. (2003). Success and failure in distance education: Perceptions of South African students and lecturers in business management. *South African Journal of Higher Education, 17, 2,* 147-157.

Koch, E., Foxcroft, C., & Watson, A. (2001). A developmental focus to student access at the university of Port Elizabeth: Process and preliminary insights in placement assessment. *South African Journal of Higher Education,15, 2*, 126-131.

Koch, S. E., & Foxcroft, C. D. (2003). A developmental focus to admissions testing: Admissions and placement standards development. *South African Journal of Higher Education, 17, 3,* 192-208.

Kotze, M., & Griessel, L. (2008). Predicting the academic performance of MBA students: A South African study. *The International Journal of the Humanities, 6*,148-156.

Koys, D. (2010). GMAT versus alternatives: Predictive validity evidence from Central Europe and the Middle East. *Journal of Education for Business, 85*, 180-185.

Lavin, D. E. (1965). The prediction of academic performance. USA: Sage Foundation.

Leedy, P. D., & Ormond, J. E. (2005). *Practical research: Planning and design* (8^{th} *ed.*). New Jersey: Upper Saddle River.

Lewis, J. (1993). *Wits Business School: Psychometric testing.* Unpublished report. University of the Witwatersrand, Johannesburg.

Lombard, R. B. (1975). *Test administrator's manual for the high level battery (A/75), revised edition.* Johannesburg: National Institute for Personnel Research, Council for Scientific and Industrial Research.

Lopes, A., Roodt, G., & Mauer, R. (2001). The predictive validity of the Apil B in a financial institution. *Journal of Industrial Psychology, 27, 1*, 61-69.

MacFarlane, M. (2006). *Predictors of academic achievement in multilingual learners.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Madaus, G. F., Scriven, M., & Stufflebeam, D. L. (1983). *Evaluation models: Viewpoints on educational and human services evaluation.* Boston: Kluwer-Nijhoff.

Magadla, L. (1994). *Academic success at university: Does the open school offer a better chance to the African black student than the government school?* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Mokone, M. (2008). *Three dimensional spatial perception and academic performance in engineering graphics: A longitudinal investigation.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Mulder, K. L. (2004). *Access to higher education and the prediction of academic success.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Murphy, K. R., & Davidshofer, C. O. (2005). *Psychological testing: Principles and applications* (*6th ed.*). New Jersey: Upper Saddle River.

Ngoako, P. (year not specified). *The high level battery.* Unpublished report. Human Sciences Research Council, Pretoria.

Nurm, J. E., & Aunola, K. (2001). How does academic achievement come about: Cross cultural and methodological notes. *International Journal of Educational Research, 35,* 403-409.

Nzimande, B. L. (1995). *To test or not to test?* Unpublished paper presented at the Congress of Psychometrics for Psychologists and Personnel Practitioners, Pretoria, June 1995.

Osborne, J.W., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research and Evaluation, 8, 2.* ISSN15317714.

Paolillo, J. G. P. (1982). The predictive validity of selected admissions variable relative to grade point average earned in a Master of Business Administration programme. *Educational and Psychological Measurement, 42,* 1163-1167.

Patterson, H., & Uys, K. (2005). Critical issues in psychological test use in the South African workplace. *South African Journal of Industrial Psychology, 31, 3,* 12-22.

Payne, J. (2008). *An examination of student factors related to performance on an undergraduate research skills course.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Peng, C. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96, 1,* 3-14.

Potter, C. S., Meyer, M. I., Scott, A. S. & Da Silva, M. (1984). Study Habits and Attitudes of Students in Five Engineering Disciplines. Johannesburg: University of the Witwatersrand. Centre for Continuing Education, Cadet Scheme Working Paper 18.33 (iv). In C.S. Potter (2004). *Measuring outcomes.* Vancouver: Commonwealth of Learning. http:\www.col.org\ PREST Handbook B2.

Potter, C. S. (1991). *An analytical case study of a pre university project.* Unpublished PhD thesis. University of the Witwatersrand, Johannesburg.

Potter, C. S. (1998). Evaluation, knowledge and human interests. *Journal of Educational Evaluation, Special Edition, 7*, 1-15.

Potter, C. S. (2004). *Measuring outcomes.* Vancouver: Commonwealth of Learning. http://www.col.org.\ PREST Handbook B2.

Potter, C. S. (2002). Programme evaluation. In M. Terre Blanche and K. Durrheim (Eds.), *Research methodology in the social sciences in Southern Africa* (2nd ed., pp. 209-226). Cape Town: University of Cape Town Press.

Potter, C. S. (in press). Extending the assessment envelope. In R. Basson, P. Fridjhon, S. Porter & C.S. Potter (Eds.), *Evaluation design and rigour in developing world contexts.* A Virtual Conference Publication of the Wits Programme Evaluation Group http://wpeg.wits.ac.za with an introduction by Stewart I. Donaldson. Johannesburg: Malagasy Press. Publication March 2012.

Ramodibe, M. R. (2008). *Organisational culture as perceived by employment equity candidates.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Randsell, S. (2001). Predicting college success: The importance of ability and non-cognitive variables. *International Journal of Educational Research, 35*, 357-364.

Ratangee, N. (2006). *Academic literacy, the PTEEP and the prediction of academic success.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Republic of South Africa (1998). Employment Equity Act (55 of 1998). Government Gazette, 400 No 19370: Pretoria (English version).

Ridgell, S. D., & Lounsbury, J. W. (2004). Predicting academic success: General intelligence, "Big Five" personality traits, and work drive. *College Student Journal, 38, 4*, 607-618.

Riordaan, M. C. (2002). *A preliminary investigation into the use of the non-cognitive questionnaire (NCO) with a sample of university applicants in South Africa.* Unpublished master's dissertation. University of Port Elizabeth, Port Elizabeth.

Rosenthal, R., & Rosnow, R. (1991). *Essentials of behavioural research: Methods and data analysis* (2nd *ed.*). USA: McGraw Hill.

Rust, J., & Golombok, S. (1992). *Modern psychometrics: The science of psychological assessment.* London: Routledge.

Saunders, E. M. (1993). *A review of the psychometric method of assessment and development interventions within an affirmative action paradigm.* Unpublished master's dissertation. University of Natal, Natal.

Sehlapelo, M., & Terre Blanche, M. (1996). Psychometric testing in South Africa. Views from above and below. *Psychology in Society, 21*, 49-59.

Shochet, I.M. (1986). *Manifest and potential performance in advantaged and disadvantaged students.* Unpublished PhD thesis. University of the Witwatersrand, Johannesburg.

Stecher, B.M., & Davis, W.K. (1987). *How to focus an evaluation.* Newbury Park, CA: Sage Publications.

Strasheim, C. (1998). *Product and service quality in management education.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Stufflebeam, D. L. (1972). The relevance of the CIPP evaluation model for educational accountability. *SRIS Quarterly, 5*, 3-6.

Tayler, J. (1996). *Assessment of the predictive validity of the learning ability battery.* Unpublished Master of Arts report. University of the Witwatersrand, Johannesburg.

Taylor, J. M. (1978). *Short term evaluation of a selection battery for trainee engineering technicians.* CSIR Special Report PERS 270. Johannesburg: National Institute of Personnel Research.

Taylor, J. M. (1980). *Factors associated with success in an engineering drawing and design course.* CSIR Special Report PERS 314. Johannesburg: National Institute of Personnel Research.

Taylor, J. M. (1981). *Factors affecting results in an engineering drawing and design course: Part two.* CSIR Special Report PERS 314. Johannesburg: National Institute of Personnel Research.

Taylor, T. R. (1987). *Test bias: The roles and responsibilities of test user and test publisher.* Pretoria: Human Sciences Research Council.

Tlakula, L. (2009). *The perceived value of international versus South African MBA programmes.* Unpublished master's dissertation. University of the Witwatersrand, Johannesburg.

Trochim, W. M. K. (2006). *Research methods: Knowledge base.* Beverly Hills: Sage.

Urbina, S. (2004). *Essentials of psychological testing.* New Jersey: John Wiley & Sons.

Viedge, C., & De Villiers, G. (1998). *Confidential feedback report from psychological assessment.* Unpublished report. Wits Business School, Johannesburg.

Visser, B. L. (1977). *Vocational counseling at the NIPR.* CSIR Special Report PERS 257. Johannesburg: National Institute of Personnel Research.

Visser, B. L. (1978). *Development of a testing programme for black university students for purposes of selection and career counseling.* NIPR. Pers 281. Johannesburg: National Institute of Personnel Research.

Visser, S. (1981). *A Short-term evaluation of the effectiveness of various NIPR tests in predicting success in engineering analysis and design.* CSIR Special Report PERS 317. Johannesburg: National Institute for Personnel Research.

Weavind, T. (2011, January 19). *What's an MBA worth today? Sunday Times.*

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, I. (1966). *Unobtrusive measures: Non-reactive research in the social sciences.* Chicago IL: Rand-McNally.

Wiggan, G. (2007). Race, school achievement, and educational inequality: Toward a student-based inquiry perspective. *Review of Educational Research, 7, 3,* 310-333.

Wits Business School (2007). *MBA prospectus.* Johannesburg: University of the Witwatersrand.

Wright, B. L. (2009). Racial-ethnic identity, academic achievement, and African American males: A review of literature. *The Journal of Negro Education, 78, 2,* 123-134.

Wright, R. E., & Palmer, J. C. (1994). GMAT scores and undergraduate GPA's as predictors of performance in graduate business programs. *Journal of Education for Business, 69, 6,* 344-348.

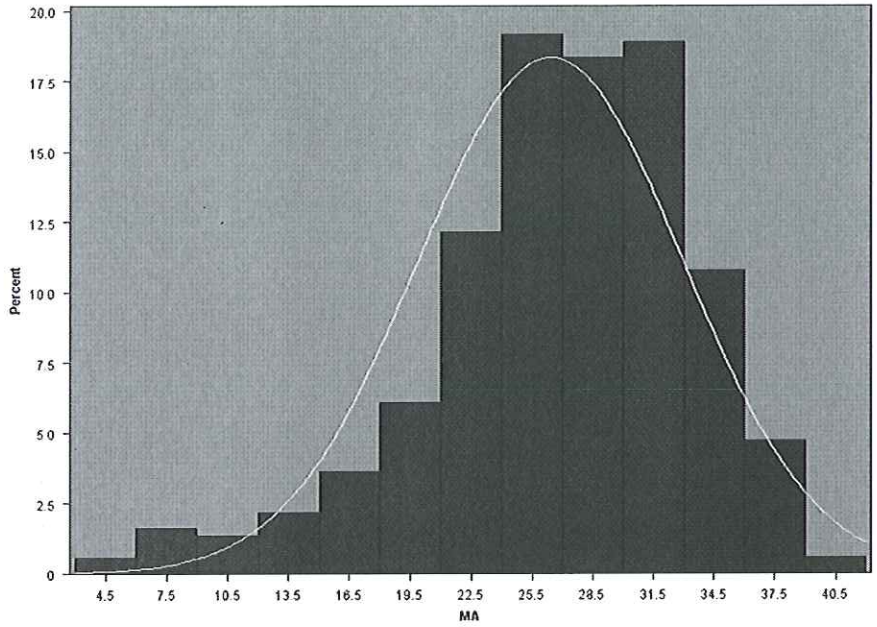
Yang, B., & Lu D. R. (2001). Predicting academic performance in management education: An empirical investigation of MBA success. *Journal of Education for Business, Sept/Oct 2001,* 15-20.

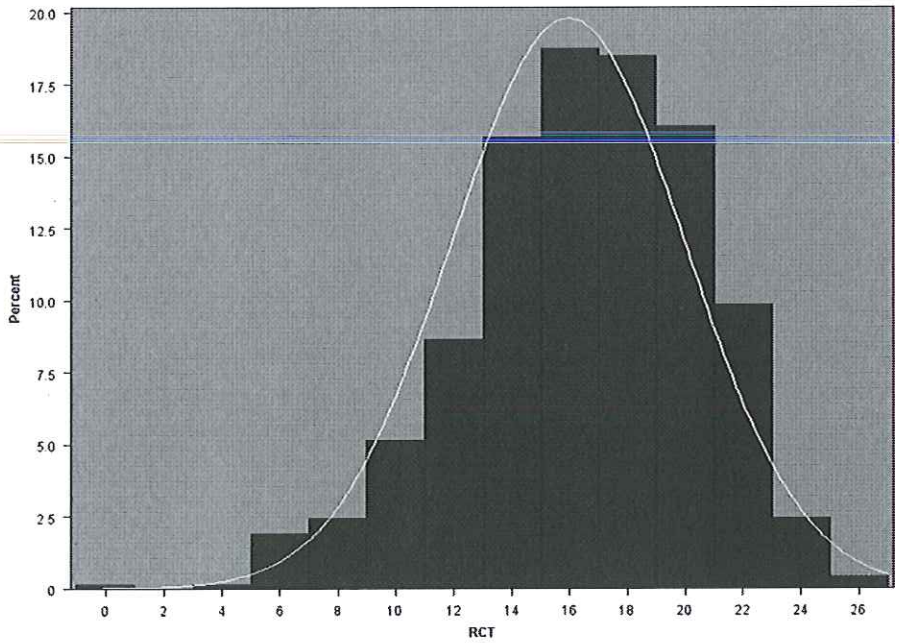
APPENDIX A: Description of Principle and Core Courses

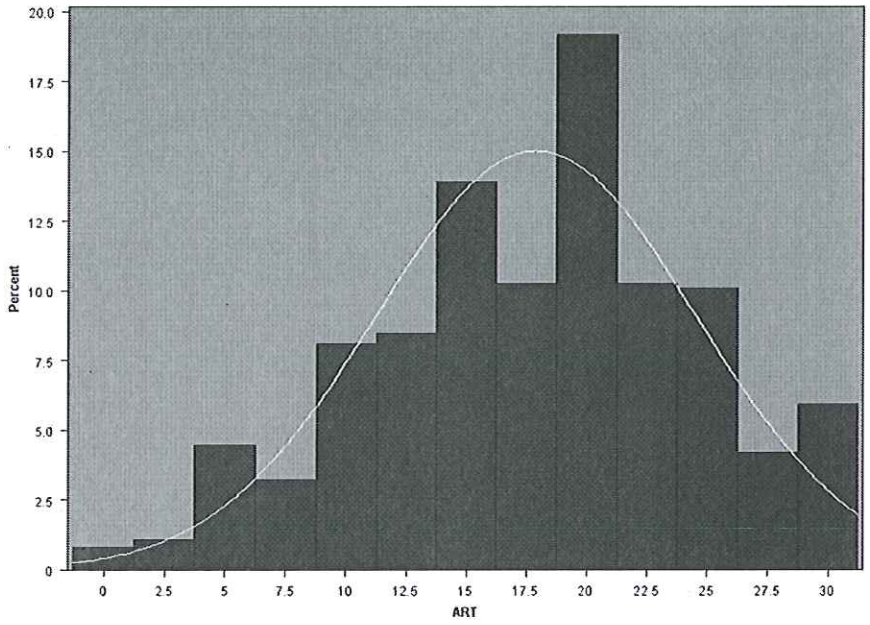
Name	Description
Principles of Economics	The course gives an overview of business and the economy. An analysis of consumer demand and production costs in competitive environments will develop an understanding of pricing strategies in the Micro economic environment.
Principles of Accounting and Finance	Assists students to develop financial literacy as generalist users of accounting information for the purpose of business decision making and control.
Principles of Human Resources	This course develops an understanding of the human aspects of business Management. It provides insight into the factors influencing attitudes in the Workplace.
Principles of Marketing	The course introduces candidates to the marketing concepts. Marketing Research and information systems, consumer behaviour, market segmentation and target market selection will be discussed.
Principles of Operations	The ability to use good judgement in any type of operating environment is important-service or manufacturing, unique through to continuous types operations. This introduction gives an environmental link to operations and a strategic framework.
Principles of Strategic Management	This course gives students a basic understanding of the role of strategy in an organisation.
Data and Decision Making (DDM)	The course is designed to broaden and deepen the student's understanding of the analytical techniques used to solve business problems in management situations.
Finance/Financial Management	Covers the management of working capital, financial analysis, the time value of money, capital budgeting, risk and return, capital structure and equity evaluation. The course intends to expose the student to the financial techniques which drive the Management of a company.
Information and Knowledge Management (IKM)	The course commences with an overview of the concepts, issues and technology relevant to information and knowledge management. Key topics cover knowledge management, electronic networks, including electronic commerce and the internet sourcing information system, the management of corporate data and information, the dynamics and trends of the information technology industry, the organisation effects of information technology, and the implementation of information systems.
Industrial Relations (IR)	The course is divided into 2 sections. The inherency of conflict between capital and labour is illustrated by describing how in South Africa those respective groups have designed structures of governance and developed processes of negotiation and dispute resolution to shape their relationship. Characteristics of the SA labour market are discussed in the second Section.
Marketing Management	The course addresses marketing planning and strategy. Special topics Include business marketing, product management and strategy, the marketing of services, distribution strategy,

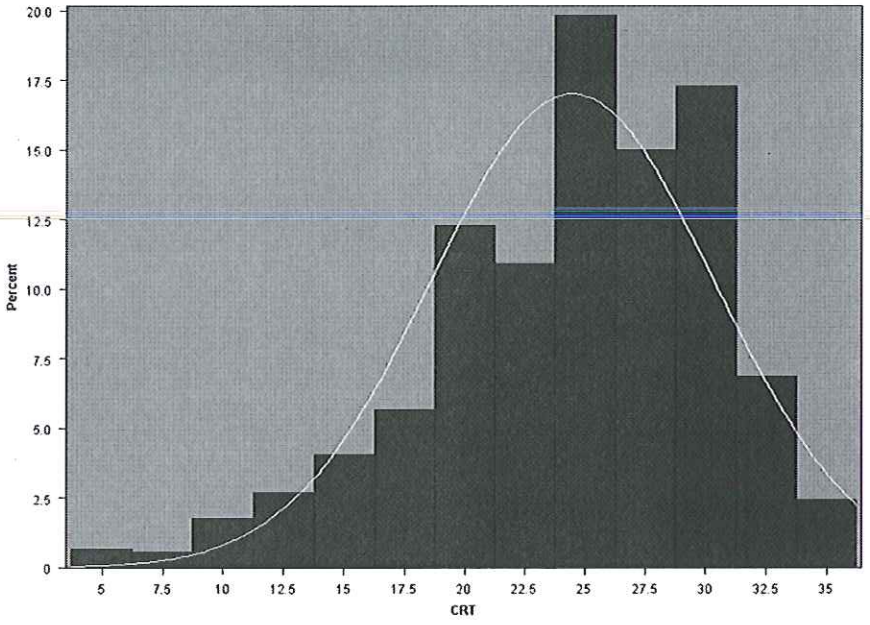
	sales management, communication management, relationship marketing and technology and retailing.
Management Accounting	The course presupposes awareness of the fundamentals of financial statements. The contents of published financial statements is now considered in greater detail. The objective is to enable students to read and understand financial statements of a company and thereby enhance the decision making ability of management.
Organisational Design and development (ODD)	The course enables students to use appropriate outcomes to analyse and critically evaluate organisational design holistically, to identify appropriate processes to institute organisational and individual change strategies, to evaluate the implications of equity and diversity management in organisations and identify processes to address issues of organisational culture, to apply the principles of and process of organisational development and process consulting to any organisational setting.
Operations Technology	The objective of the course is to enable students to build simple but correctly formulated models of any operation such that useful pictures and or conclusions can be drawn from the process.
Strategic Management	The course gives students an understanding of thinking approaches and actions that Allow for the creation and execution of strategic change in an organisation. It will Identify local and global strategic leadership and management issues.
Research Methodology	Students are briefly exposed to the philosophy of science and an understanding of the
	scientific method. Qualitative and quantitative approaches to research are contrasted and the methods of research design and analysis are explored for both paradigms. Students are given guidance to the writing of their research proposals and their research reports.

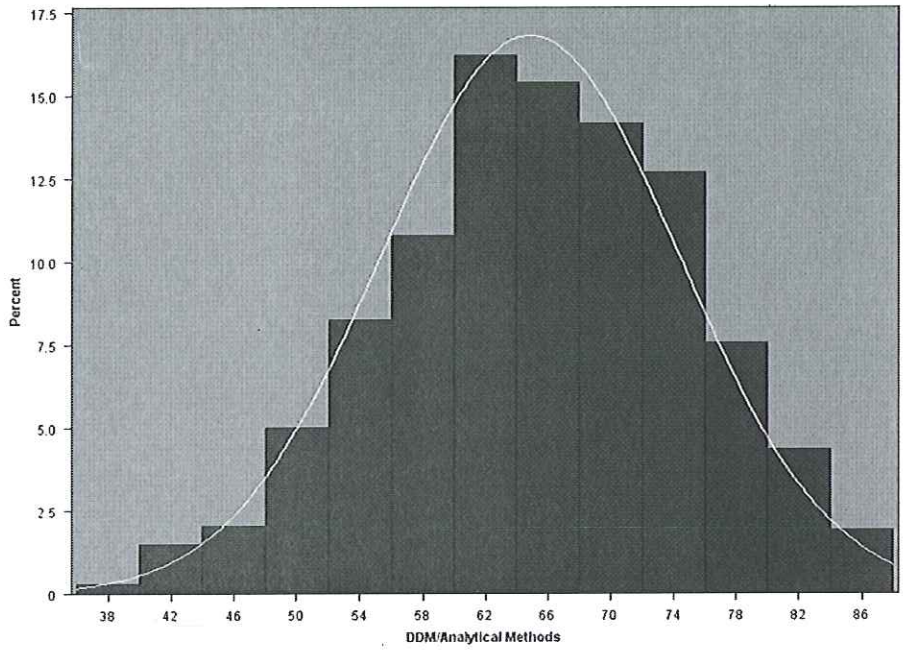
Appendix B: Normal Distribution Diagrams

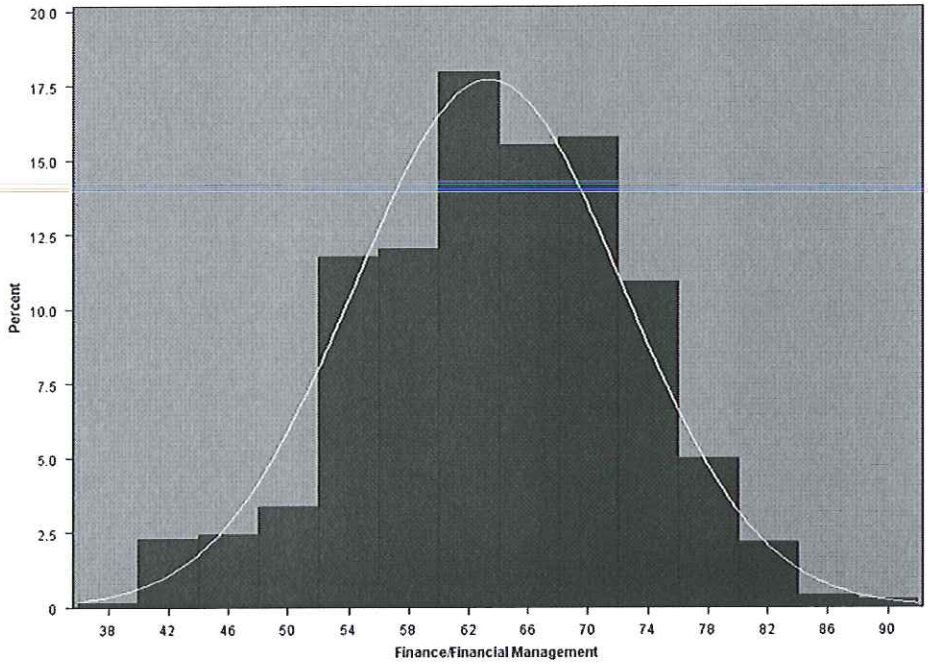


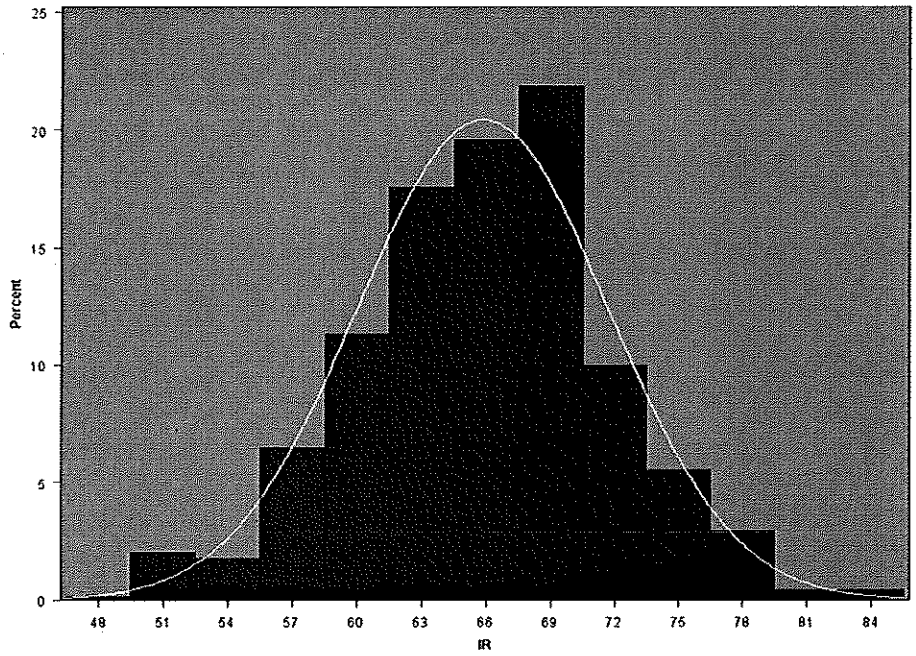


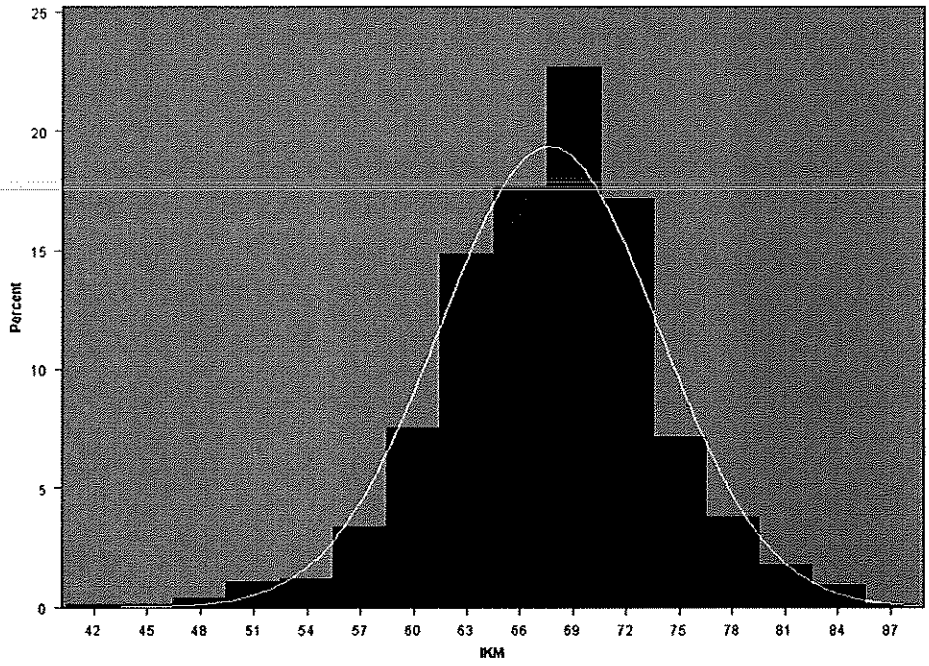

Percent
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
0
2.5
5
7.5
10
12.5
15
17.5
20
22.5
25
27.5
30
ART

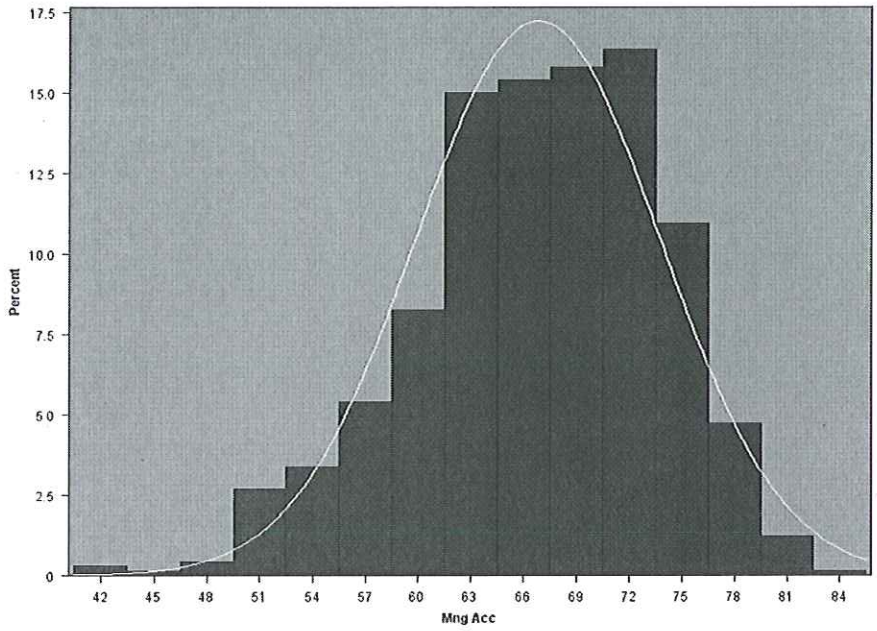

Percent
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
5
7.5
10
12.5
15
17.5
20
22.5
25
27.5
30
32.5
35
CRT

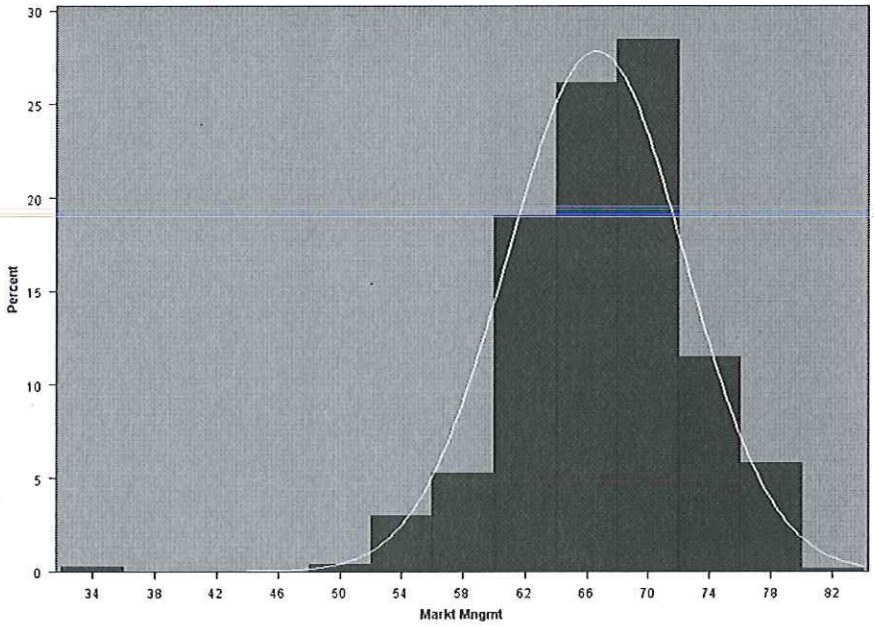

Percent
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
38
42
46
50
54
58
62
66
70
74
78
82
86
DDM/Analytical Methods

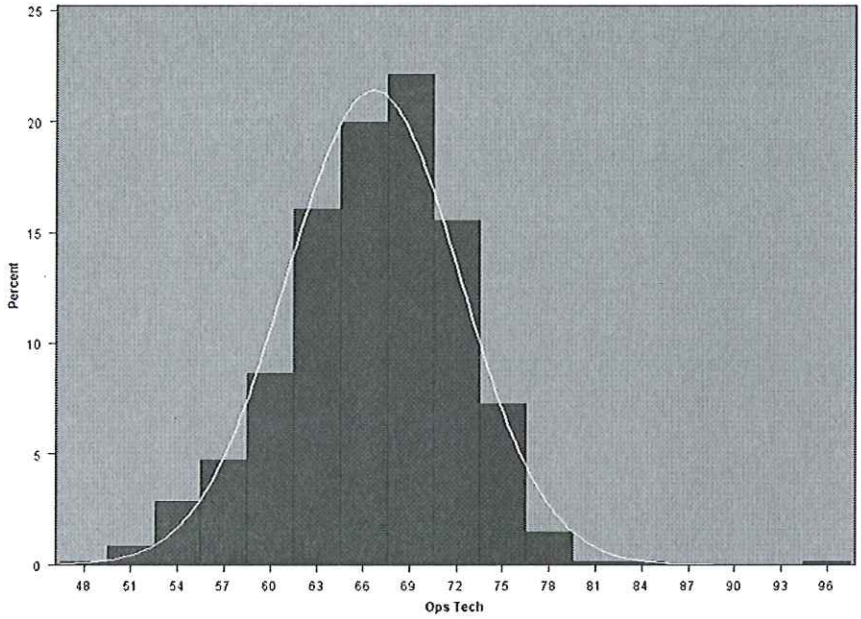

Percent
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
38
42
46
50
54
58
62
66
70
74
78
82
86
90
Finance/Financial Management

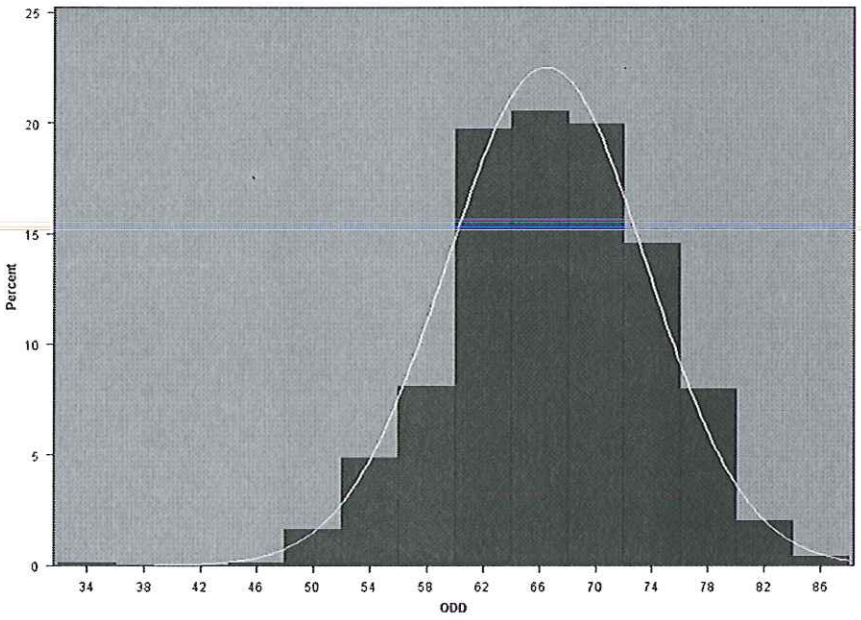

Percent
0
5
10
15
20
25
48
51
54
57
60
63
66
69
72
75
78
81
84
IR

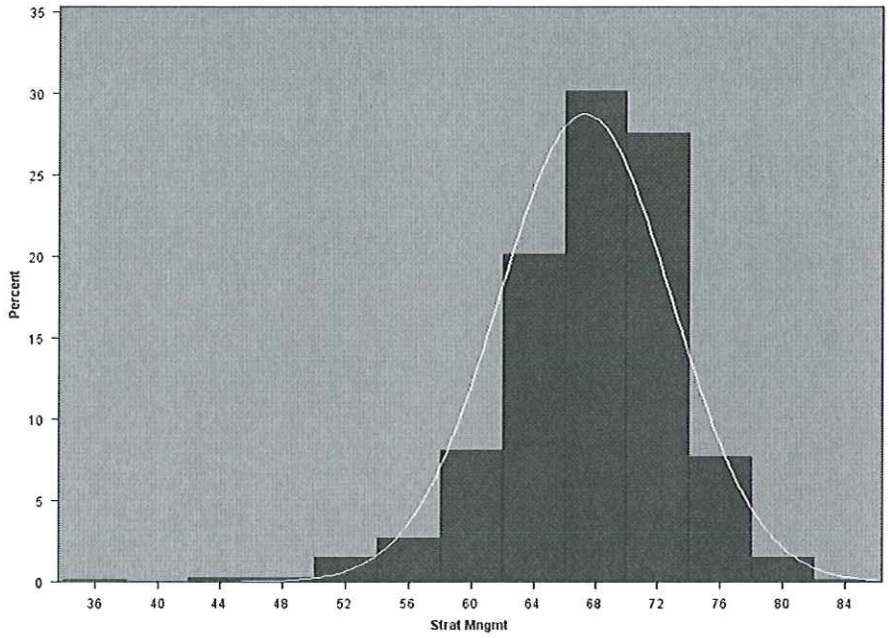

Percent
0
5
10
15
20
25
42
45
48
51
54
57
60
63
66
69
72
75
78
81
84
87
IKM

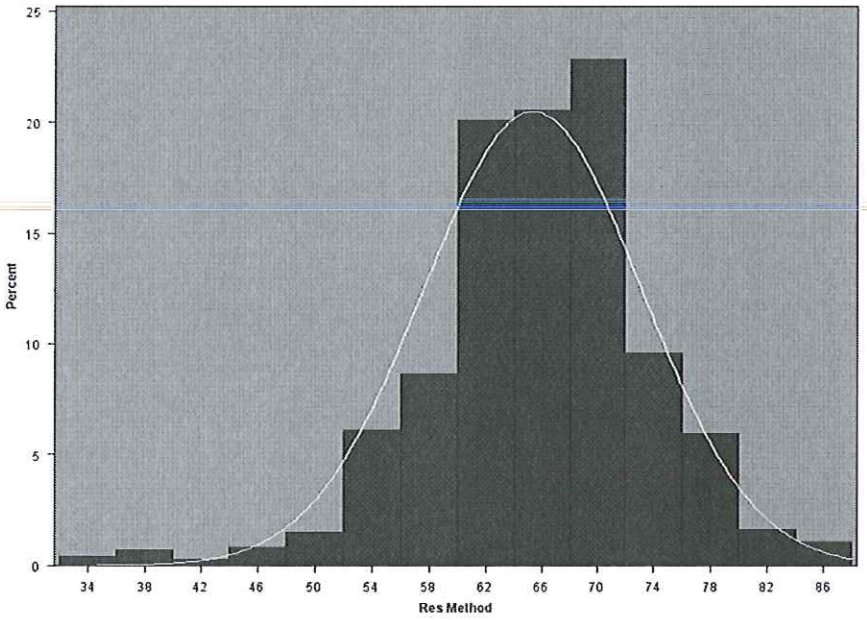

Percent
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
42
45
48
51
54
57
60
63
66
69
72
75
78
81
84
Mng Acc

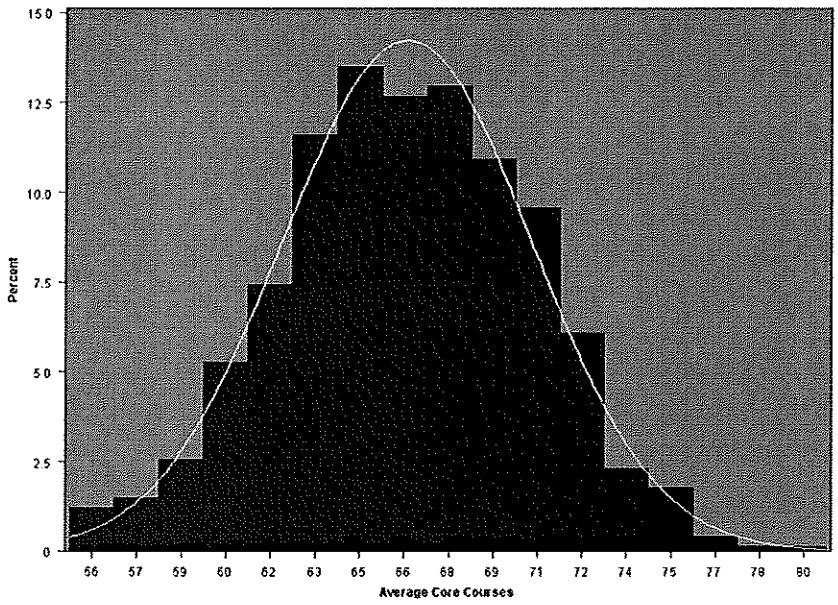

Percent
30
25
20
15
10
5
0
34
38
42
46
50
54
58
62
66
70
74
78
82
Markt Mngmt


Percent
25
20
15
10
5
0
48
51
54
57
60
63
66
69
72
75
78
81
84
87
90
93
96
Ops Tech

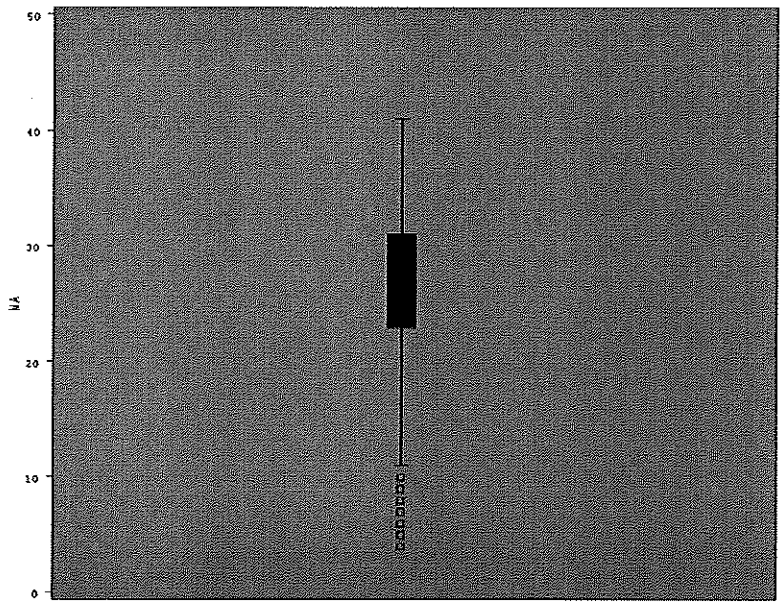

Percent
25
20
15
10
5
0
34
38
42
46
50
54
58
62
66
70
74
78
82
86
ODD

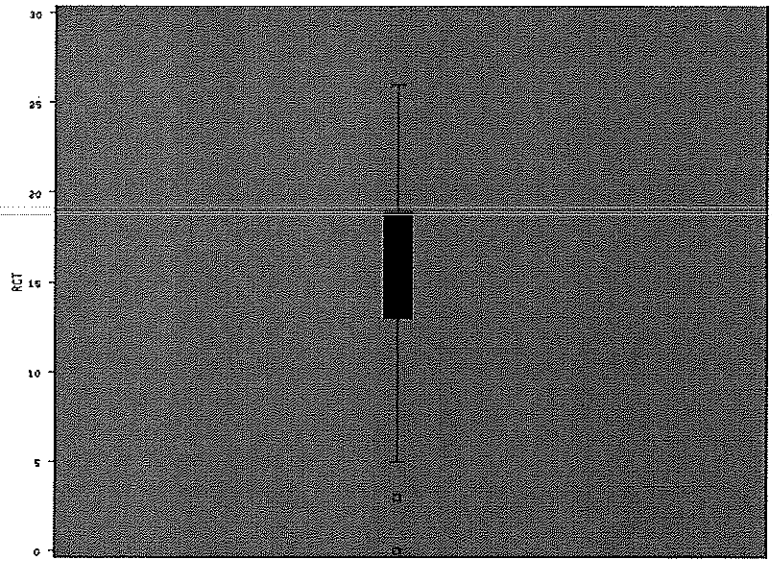

Percent
35
30
25
20
15
10
5
0
36
40
44
48
52
56
60
64
68
72
76
80
84
Strat Mngmt

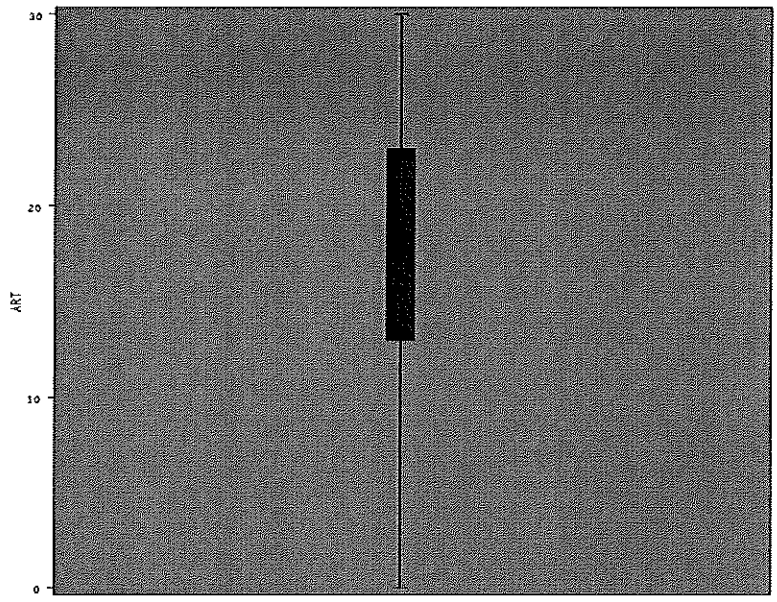

Percent
25
20
15
10
5
0
34
38
42
46
50
54
58
62
66
70
74
78
82
86
Res Method

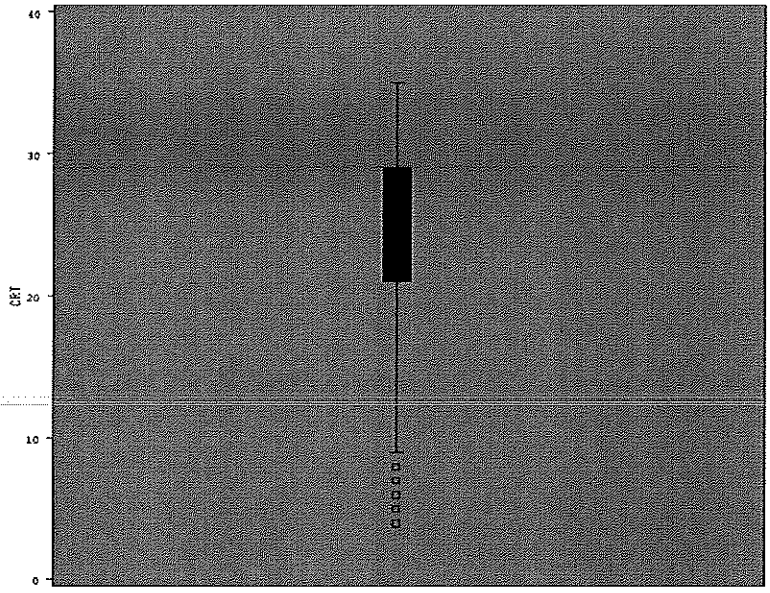

15.0
12.5
10.0
7.5
5.0
2.5
0
Percent
56
57
59
60
62
63
65
66
68
69
71
72
74
75
77
78
80
Average Core Courses

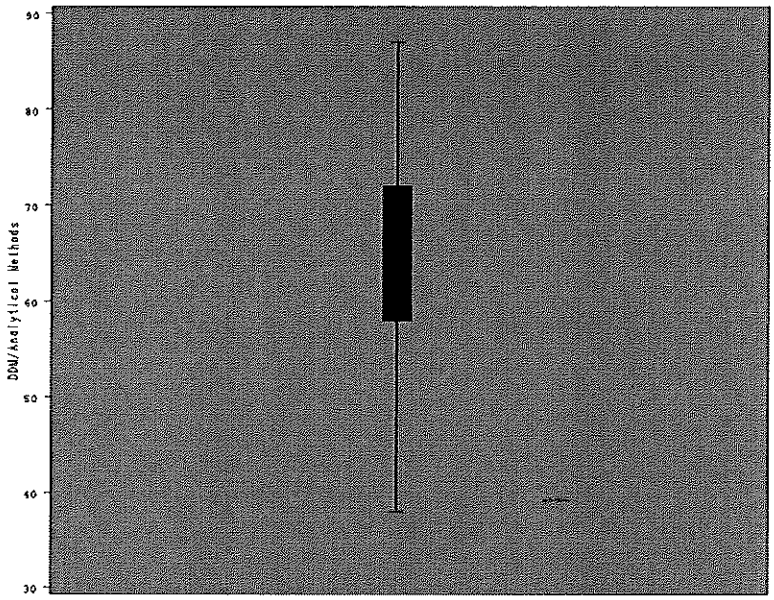
APPENDIX C: BOX-WHISKER PLOTS

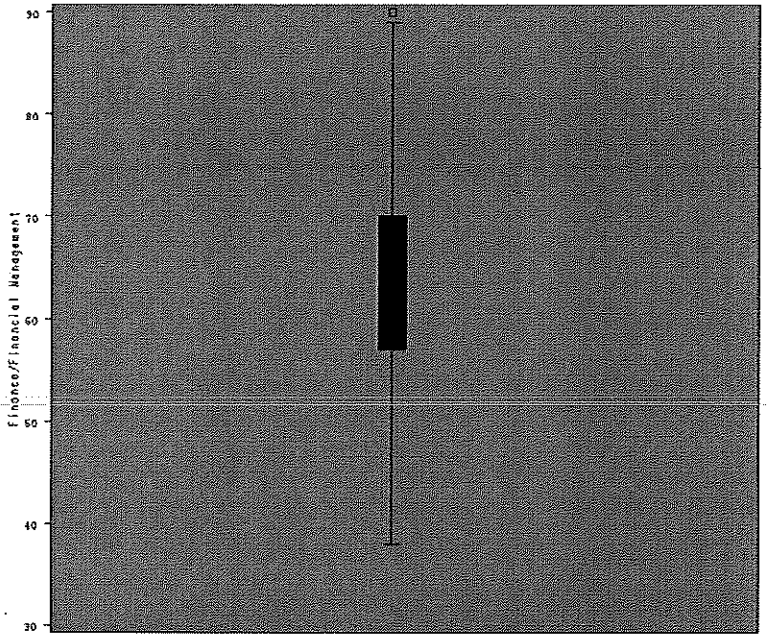


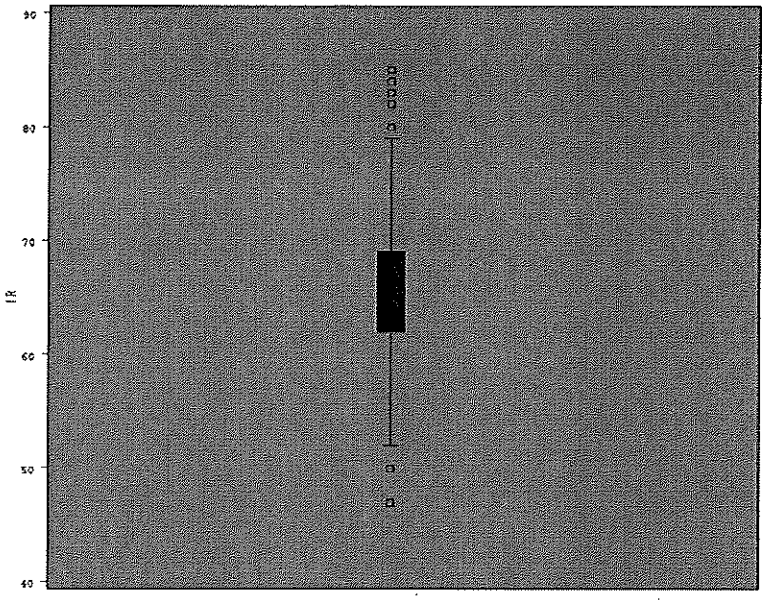


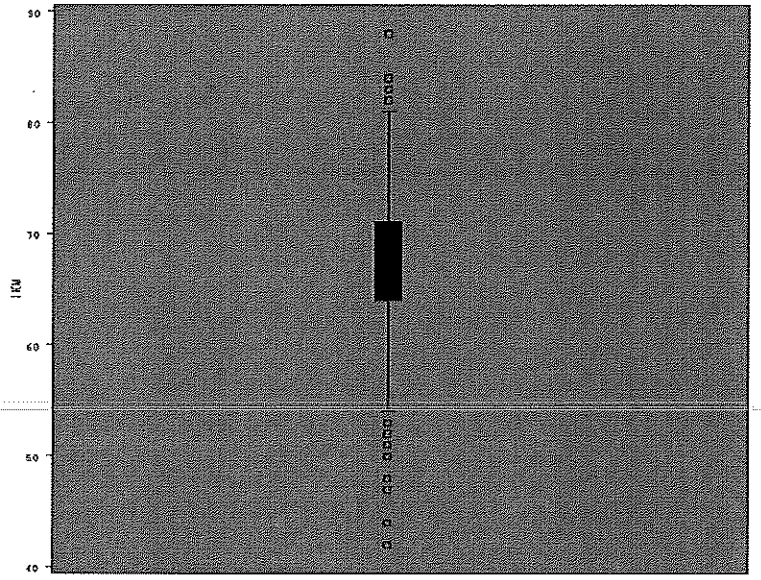

30
20
10
0
ART

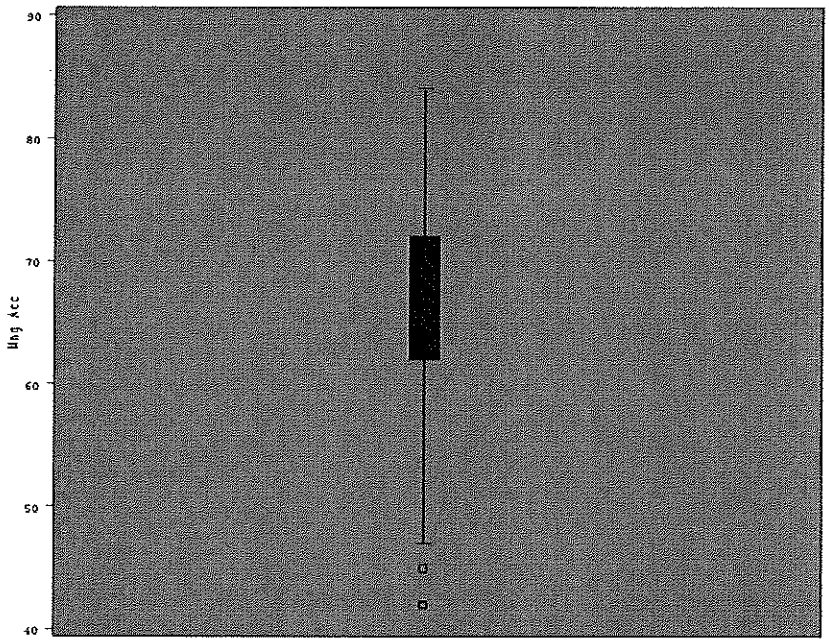

40
30
20
10
0
CRT

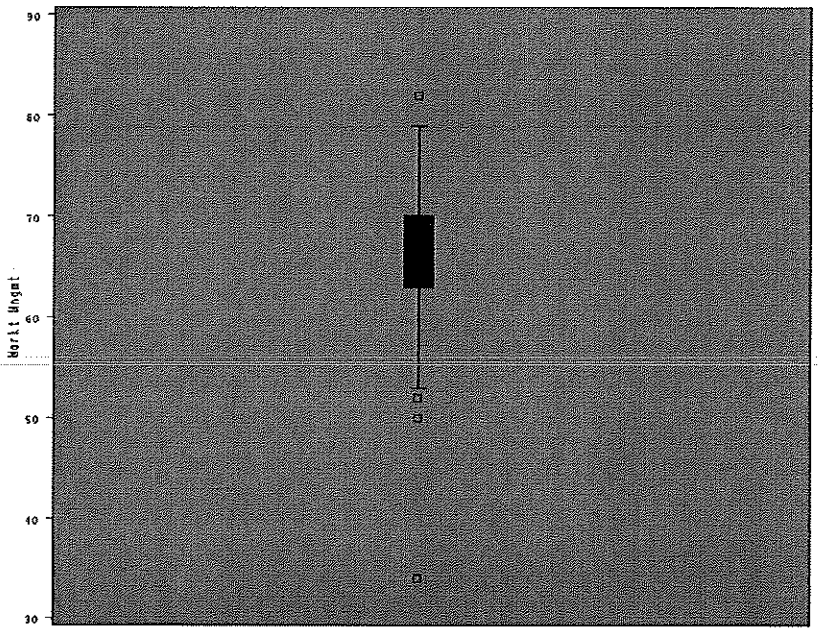

90
80
70
60
50
40
30
DDM/Analytical Methods

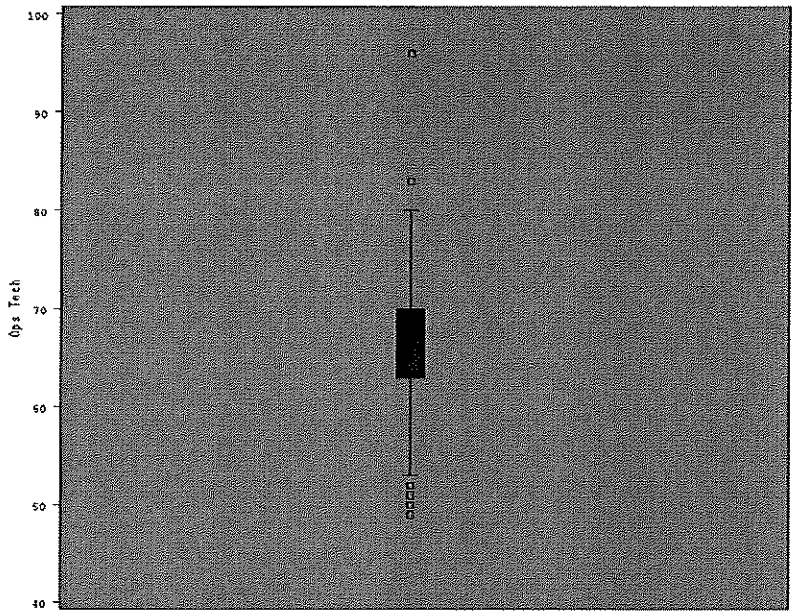

90
80
70
60
50
40
30
Finance/Financial Management

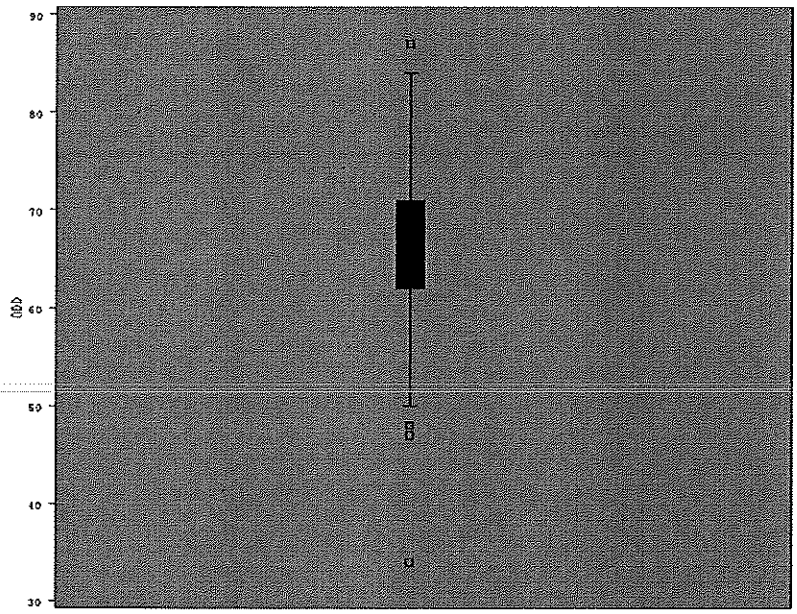

90
80
70
60
50
40
IR

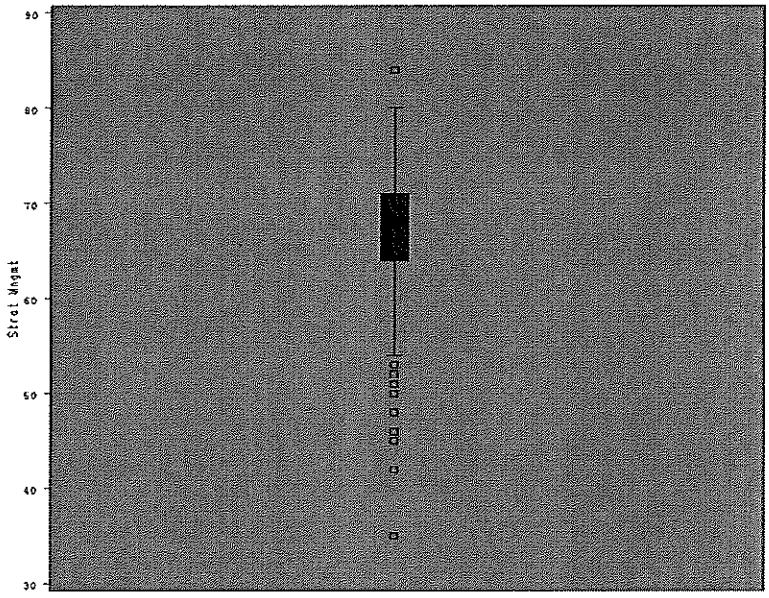

90
80
70
60
50
40
IRW

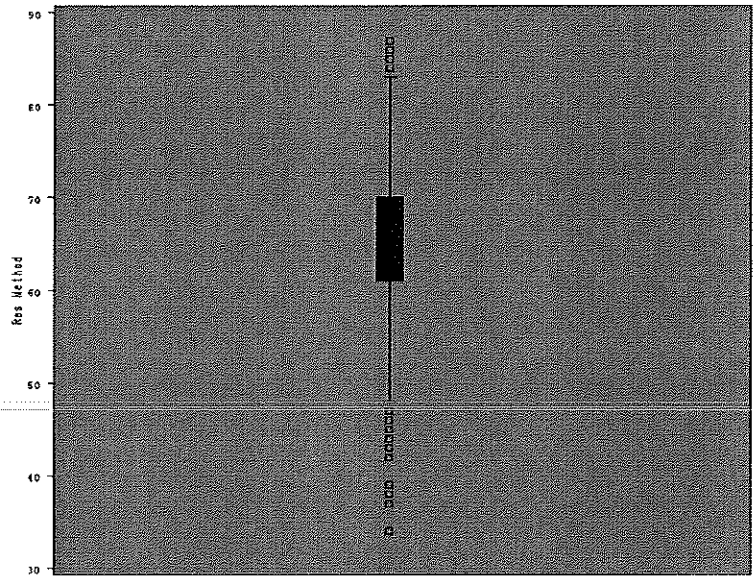

90
80
70
60
50
40
Ung Acc

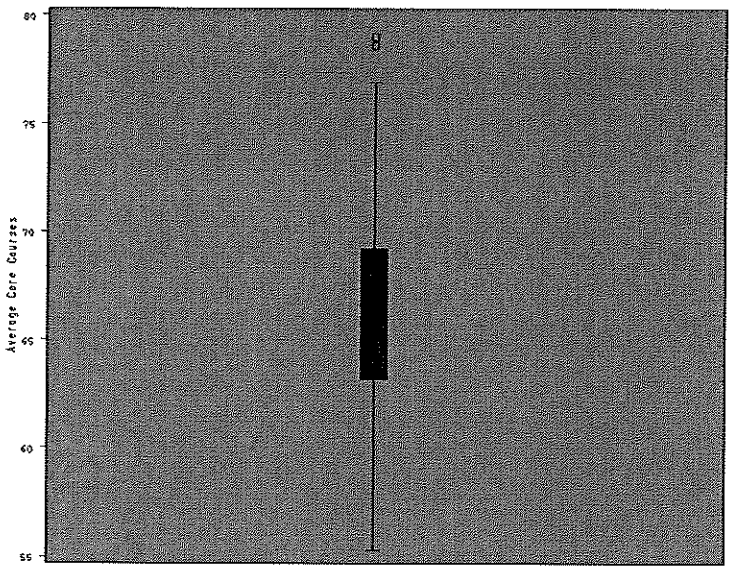

90
80
70
60
50
40
30
Markt Ungat


100
90
80
70
60
50
40
Ops Tech


90
80
70
60
50
40
30
QQD


Strat Nagat
90
80
70
60
50
40
30

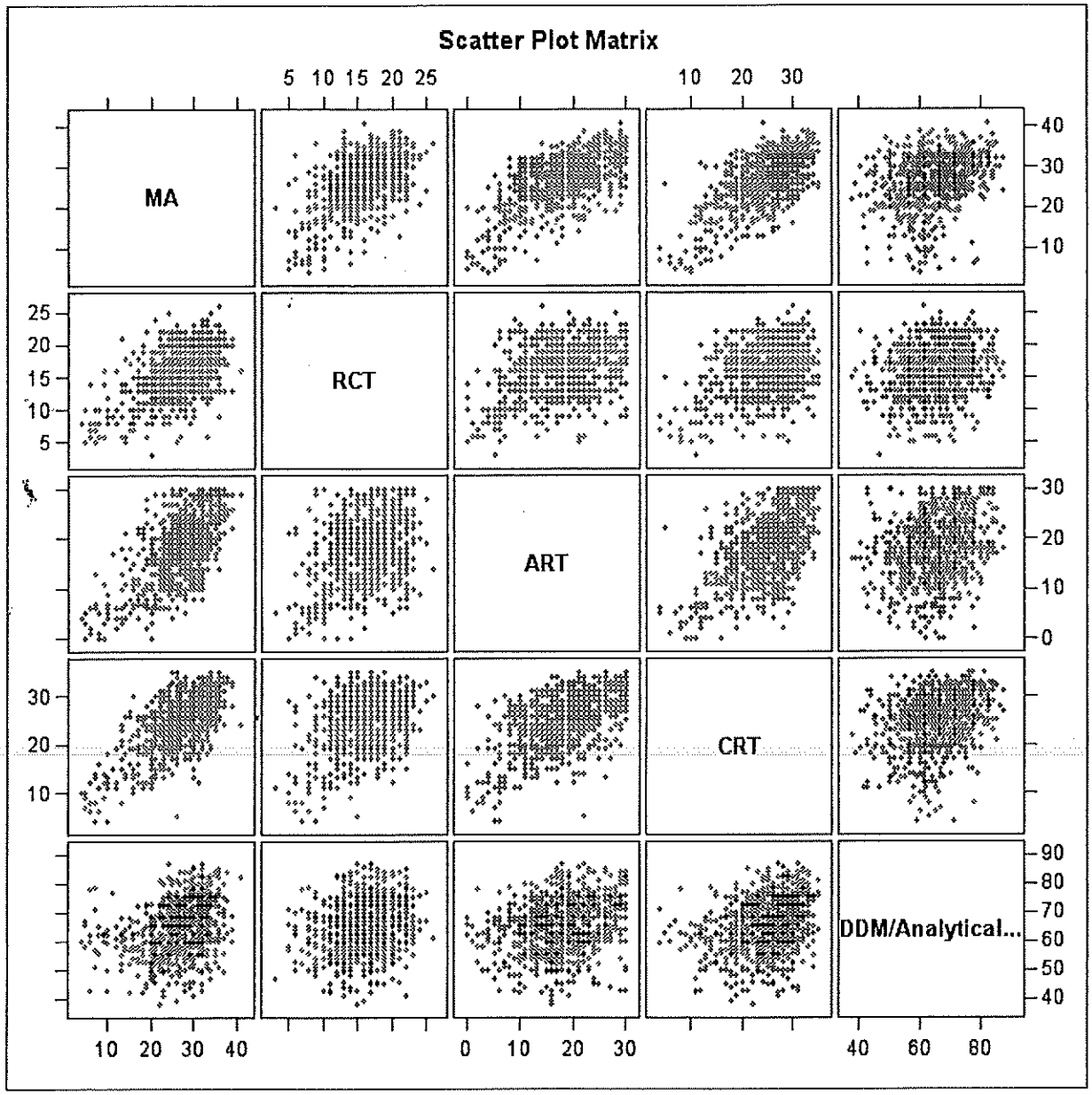

Res Method
90
80
70
60
50
40
30


80
75
70
65
60
55
Average Core Courses
8

Appendix D: Correlation Matrix

Pearson Correlation Coefficients, N = 741 Prob > |r| under H0: Rho=0															
	MA	RCT	ART	CRT	DDM/Analytical Methods	Finance/Financial Management	IR	IKM	Mng Acc	Markt Mngmt	Ops Tech	ODD	Strat Mngmt	Res Method	Average Core Courses
MA	1	0.5384	0.63248	0.66512	0.29942	0.31775	0.06902	0.10826	0.26983	0.1004	0.28889	0.11457	0.1718	0.09768	0.31833
		<.0001	<.0001	<.0001	<.0001	<.0001	0.0604	0.0032	<.0001	0.0062	<.0001	0.0018	<.0001	0.0078	<.0001
RCT	0.5384	1	0.32586	0.39513	0.20418	0.19897	0.13558	0.14682	0.25127	0.05849	0.1967	0.14132	0.14765	0.10516	0.26793
	<.0001		<.0001	<.0001	<.0001	<.0001	0.0002	<.0001	<.0001	0.1116	<.0001	0.0001	<.0001	0.0042	<.0001
ART	0.63248	0.32586	1	0.56717	0.26012	0.29393	0.03769	0.04083	0.14292	0.08302	0.25409	0.03469	0.09153	0.05408	0.22988
	<.0001	<.0001		<.0001	<.0001	<.0001	0.3055	0.267	<.0001	0.0238	<.0001	0.3456	0.0127	0.1413	<.0001
CRT	0.66512	0.39513	0.56717	1	0.29585	0.27929	0.04895	0.09697	0.15972	0.06908	0.30744	0.06503	0.14894	0.09376	0.27269
	<.0001	<.0001	<.0001		<.0001	<.0001	0.1832	0.0083	<.0001	0.0602	<.0001	0.0769	<.0001	0.0107	<.0001
DDM/Analy	0.29942	0.20418	0.26012	0.29585	1	0.52389	0.26746	0.21912	0.2853	0.30054	0.40871	0.26684	0.21978	0.2921	0.67876
	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Finance/Fi	0.31775	0.19897	0.29393	0.27929	0.52389	1	0.17353	0.18951	0.37528	0.23314	0.4291	0.16095	0.21252	0.2503	0.63768
	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
IR	0.06902	0.13558	0.03769	0.04895	0.26746	0.17353	1	0.33094	0.24026	0.25165	0.2038	0.28083	0.21669	0.29412	0.51759
	0.0604	0.0002	0.3055	0.1832	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
IKM	0.10826	0.14682	0.04083	0.09697	0.21912	0.18951	0.33094	1	0.29022	0.2602	0.28497	0.32606	0.24458	0.27084	0.54258
	0.0032	<.0001	0.267	0.0083	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Mng Acc	0.26983	0.25127	0.14292	0.15972	0.2853	0.37528	0.24026	0.29022	1	0.30109	0.39995	0.26561	0.31429	0.25232	0.61409
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Markt Mng	0.1004	0.05849	0.08302	0.06908	0.30054	0.23314	0.25165	0.2602	0.30109	1	0.39726	0.33827	0.27976	0.36581	0.59234
	0.0062	0.1116	0.0238	0.0602	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001
Ops Tech	0.28889	0.1967	0.25409	0.30744	0.40871	0.4291	0.2038	0.28497	0.39995	0.39726	1	0.30869	0.39478	0.33171	0.67404
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001
ODD	0.11457	0.14132	0.03469	0.06503	0.26684	0.16095	0.28083	0.32606	0.26561	0.33827	0.30869	1	0.29512	0.34865	0.58551
	0.0018	0.0001	0.3456	0.0769	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001
Strat Mngr	0.1718	0.14765	0.09153	0.14894	0.21978	0.21252	0.21669	0.24458	0.31429	0.27976	0.39478	0.29512	1	0.30414	0.5426
	<.0001	<.0001	0.0127	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001
Res Metho	0.09768	0.10516	0.05408	0.09376	0.2921	0.2503	0.29412	0.27084	0.25232	0.36581	0.33171	0.34865	0.30414	1	0.62086
	0.0078	0.0042	0.1413	0.0107	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001
Average C	0.31833	0.26793	0.22988	0.27269	0.67876	0.63768	0.51759	0.54258	0.61409	0.59234	0.67404	0.58551	0.5426	0.62086	1
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	

APPENDIX E: SCATTERPLOTS OF TEST SCORES



Appendix F: Multiple Regression

Dependent Variable: Average Core Courses							
Number of Observ	741						
Number of Observ	741						
Forward Selection: Step 1							
Variable MA Entered: R-Square = 0.1013 and C(p) = 15.0643							
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	1	1323.663	1323.663	83.33	<.0001		
Error	739	11739	15.88458				
Corrected Total	740	13062					
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F		
Intercept	60.67052	0.61122	156510	9852.94	<.0001		
MA	0.20499	0.02246	1323.663	83.33	<.0001		
Bounds on condition number: 1, 1							
Forward Selection: Step 2							
Variable RCT Entered: R-Square = 0.1145 and C(p) = 6.0797							
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	2	1495.118	747.5589	47.69	<.0001		
Error	738	11567	15.67378				
Corrected Total	740	13062					
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F		
Intercept	59.65534	0.68032	120515	7688.95	<.0001		
MA	0.15785	0.02647	557.382	35.56	<.0001		
RCT	0.1417	0.04284	171.4549	10.94	0.001		
Bounds on condition number: 1.4082, 5.6328							
Forward Selection: Step 3							
Variable CRT Entered: R-Square = 0.1201 and C(p) = 3.3686							
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	3	1568.652	522.8841	33.53	<.0001		
Error	737	11494	15.59527				
Corrected Total	740	13062					
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F		

Intercept	59.06884	0.73039	101999	6540.4	<.0001		
MA	0.11656	0.03254	200.1369	12.83	0.0004		
RCT	0.13623	0.04281	157.9237	10.13	0.0015		
CRT	0.07215	0.03323	73.53449	4.72	0.0302		

Bounds on condition number: 2.1385, 16.054

No other variable met the 0.5000 significance level for entry into the model.

Summary of Forward Selection

Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	MA	1	0.1013	0.1013	15.0643	83.33	<.0001
2	RCT	2	0.0131	0.1145	6.0797	10.94	0.001
3	CRT	3	0.0056	0.1201	3.3686	4.72	0.0302

Dependent Variable: Average Core Courses

Number of Observ	741						
Number of Observ	741						

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	3	1568.652	522.8841	33.53	<.0001		
Error	737	11494	15.59527				
Corrected Total	740	13062					
Root MSE	3.94908	R-Square	0.1201				
Dependent Mean	66.08758	Adj R-Sq	0.1165				
Coeff Var	5.97553						

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate	
Intercept	1	59.06884	0.73039	80.87	<.0001	0	
MA	1	0.11656	0.03254	3.58	0.0004	0.18101	
RCT	1	0.13623	0.04281	3.18	0.0015	0.13071	
CRT	1	0.07215	0.03323	2.17	0.0302	0.10065	

Correlation of Estimates

Variable	Intercept	MA	RCT	CRT			
Intercept	1	-0.1513	-0.3967	-0.3698			
MA	-0.1513	1	-0.4018	-0.5844			
RCT	-0.3967	-0.4018	1	-0.0588			
CRT	-0.3698	-0.5844	-0.0588	1			

Collinearity Diagnostics

Number	Eigenvalue	Condition Index	Proportion of Variation				
			Intercept	MA	RCT	CRT	
1	3.91632	1	0.00251	0.00172	0.0027	0.00193	
2	0.03602	10.42715	0.01327	0.05551	0.66852	0.28906	
3	0.03062	11.30871	0.90972	0.14986	0.15982	0.00729	
4	0.01703	15.16356	0.07449	0.79291	0.16897	0.70172	

Dependent Variable: Average Core Courses							
Test of First and Second							
Moment Specification							
DF	Chi-Square	Pr > ChiSq					
9	3.97	0.9134					
Durbin-Watson D	2.056						
Number of Observ	741						
1st Order Autocorr	-0.029						

APPENDIX F1: Multiple Regression Using Language Variable

'Prefered Language'n=1

Number of Observations Read	492
Number of Observations Used	492

Forward Selection: Step 1

Variable MA Entered: R-Square = 0.0874 and C(p) = 13.1899

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	717.0749	717.0749	46.96	<.0001
Error	490	7482.958	15.27134		
Corrected Total	491	8200.033			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	61.56775	0.7592	100433	6576.58	<.0001
MA	0.18756	0.02737	717.0749	46.96	<.0001

Bounds on condition number: 1, 1

Forward Selection: Step 2

Variable RCT Entered: R-Square = 0.1048 and C(p) = 5.6806

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	859.0533	429.5267	28.61	<.0001
Error	489	7340.98	15.01223		
Corrected Total	491	8200.033			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	60.53207	0.82463	80891	5388.32	<.0001
MA	0.13441	0.03217	262.0181	17.45	<.0001
RCT	0.15228	0.04952	141.9784	9.46	0.0022

Bounds on condition number: 1.4055, 5.6221

Forward Selection: Step 3

Variable CRT Entered: R-Square = 0.1115 and C(p) = 3.9727

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	914.4131	304.8044	20.42	<.0001
Error	488	7285.62	14.92955		
Corrected Total	491	8200.033			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	59.919	0.88183	68929	4616.97	<.0001
MA	0.09105	0.0392	80.55012	5.4	0.0206
RCT	0.14481	0.04953	127.6099	8.55	0.0036

CRT	0.07652	0.03974	55.35979	3.71	0.0547

Bounds on condition number: 2.0979, 15.889

Forward Selection: Step 4

Variable ART Entered: R-Square = 0.1133 and C(p) = 5.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	928.9357	232.2339	15.55	<.0001
Error	487	7271.097	14.93038		
Corrected Total	491	8200.033			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	59.87695	0.88289	68672	4599.48	<.0001
MA	0.10644	0.04219	95.02337	6.36	0.012
RCT	0.1443	0.04954	126.6919	8.49	0.0037
ART	-0.03393	0.0344	14.52262	0.97	0.3245
CRT	0.08704	0.04114	66.81533	4.48	0.0349

Bounds on condition number: 2.4301, 29.92

All variables have been entered into the model.

Summary of Forward Selection							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	MA	1	0.0874	0.0874	13.1899	46.96	<.0001
2	RCT	2	0.0173	0.1048	5.6806	9.46	0.0022
3	CRT	3	0.0068	0.1115	3.9727	3.71	0.0547
4	ART	4	0.0018	0.1133	5	0.97	0.3245

'Prefered Language'n=1

Dependent Variable: Average Core Courses

'Prefered Language'n=1

Number of Observations Read	492
Number of Observations Used	492

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	928.9357	232.2339	15.55	<.0001
Error	487	7271.097	14.93038		
Corrected Total	491	8200.033			
Root MSE	3.86399	R-Square	0.1133		
Dependent Mean	66.62805	Adj R-Sq	0.106		
Coeff Var	5.79934				

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	59.87695	0.88289	67.82	<.0001	0
MA	1	0.10644	0.04219	2.52	0.012	0.16781
RCT	1	0.1443	0.04954	2.91	0.0037	0.14782
ART	1	-0.03393	0.0344	-0.99	0.3245	-0.05523
CRT	1	0.08704	0.04114	2.12	0.0349	0.12484

Correlation of Estimates					
Variable	Intercept	MA	RCT	ART	CRT
Intercept	1	-0.2005	-0.3505	0.0483	-0.3608
MA	-0.2005	1	-0.3693	-0.3698	-0.4196
RCT	-0.3505	-0.3693	1	0.0105	-0.0783
ART	0.0483	-0.3698	0.0105	1	-0.2593
CRT	-0.3608	-0.4196	-0.0783	-0.2593	1

Collinearity Diagnostics							
			Proportion of Variation				
Number	Eigenvalue	Condition Index	Intercept	MA	RCT	ART	CRT
1	4.85555	1	0.0016	0.000926	0.00181	0.00268	0.00114
2	0.06906	8.38494	0.08595	7.36E-06	0.11139	0.64868	3.06E-05
3	0.034	11.94946	0.28289	1.5E-05	0.73994	0.04044	0.1303
4	0.02492	13.95979	0.61808	0.14365	0.00373	0.27167	0.37334
5	0.01646	17.17383	0.01148	0.8554	0.14313	0.03652	0.49518

'Prefered Language'n=1

Test of First and Second Moment Specification		
DF	Chi-Square	Pr > ChiSq
14	10.48	0.7264
Durbin-Watson D	2.083	
Number of Observations	492	
1st Order Autocorrelation	-0.042	
Number of Observations Read	249	
Number of Observations Used	249	

Forward Selection: Step 1

Variable MA Entered: R-Square = 0.1020 and C(p) = 5.3640

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	452.1395	452.1395	28.04	<.0001
Error	247	3982.514	16.12354		
Corrected Total	248	4434.654			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	59.80461	1.01716	55738	3456.95	<.0001
MA	0.20586	0.03887	452.1395	28.04	<.0001

Bounds on condition number: 1, 1

Forward Selection: Step 2

Variable ART Entered: R-Square = 0.1170 and C(p) = 3.1821

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	518.66	259.33	16.29	<.0001
Error	246	3915.994	15.91867		
Corrected Total	248	4434.654			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	59.76486	1.01086	55644	3495.49	<.0001
MA	0.13792	0.05096	116.617	7.33	0.0073
ART	0.10532	0.05152	66.52057	4.18	0.042

Bounds on condition number: 1.7403, 6.9612

Forward Selection: Step 3

Variable RCT Entered: R-Square = 0.1243 and C(p) = 3.1263

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	551.3609	183.787	11.6	<.0001
Error	245	3883.293	15.85017		
Corrected Total	248	4434.654			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	58.82584	1.20201	37962	2395.07	<.0001
MA	0.098	0.05795	45.33167	2.86	0.0921
RCT	0.11771	0.08195	32.70088	2.06	0.1522
ART	0.11329	0.05171	76.08201	4.8	0.0294

Bounds on condition number: 2.2603, 16.282

No other variable met the 0.5000 significance level for entry into the model.

Summary of Forward Selection							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	MA	1	0.102	0.102	5.364	28.04	<.0001
2	ART	2	0.015	0.117	3.1821	4.18	0.042
3	RCT	3	0.0074	0.1243	3.1263	2.06	0.1522

'Prefered Language'n=2

Number of Observations Read	249
Number of Observations Used	249

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	551.3609	183.787	11.6	<.0001
Error	245	3883.293	15.85017		
Corrected Total	248	4434.654			
Root MSE	3.98123	R-Square	0.1243		
Dependent Mean	65.01968	Adj R-Sq	0.1136		
Coeff Var	6.12311				

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	58.82584	1.20201	48.94	<.0001	0
MA	1	0.098	0.05795	1.69	0.0921	0.152
RCT	1	0.11771	0.08195	1.44	0.1522	0.10183

ART	1	0.11329	0.05171	2.19	0.0294	0.1738

Correlation of Estimates

Variable	Intercept	MA	RCT	ART
Intercept	1	-0.2702	-0.5439	-0.0744
MA	-0.2702	1	-0.4796	-0.6205
RCT	-0.5439	-0.4796	1	0.1073
ART	-0.0744	-0.6205	0.1073	1

Collinearity Diagnostics

Number	Eigenvalue	Condition Index	Proportion of Variation			
			Intercept	MA	RCT	ART
1	3.86964	1	0.00285	0.00182	0.00244	0.00467
2	0.08252	6.84769	0.08972	0.00103	0.08634	0.57411
3	0.02828	11.69842	0.90215	0.12133	0.37463	0.04156
4	0.01956	14.06494	0.00527	0.87583	0.53659	0.37965

'Prefered Language'n=2

Test of First and Second Moment Specification

DF	Chi-Square	Pr > ChiSq
9	9.15	0.4237
Durbin-Watson D	1.914	
Number of Observations	249	
1st Order Autocorrelation	0.038	

Appendix F2:Multiple Regression by Race Variable

Dependent Variable: Average Core Courses

Race=1

Number of Observations R	110
Number of Observations U	110

Forward Selection: Step 1

Variable MA Entered: R-Square = 0.0343 and C(p) = 0.8880

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	45.48877	45.48877	3.84	0.0528
Error	108	1280.982	11.86095		
Corrected Total	109	1326.471			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	63.30071	1.61176	18295	1542.48	<.0001
MA	0.11453	0.05848	45.48877	3.84	0.0528

Bounds on condition number: 1, 1

Forward Selection: Step 2

Variable ART Entered: R-Square = 0.0441 and C(p) = 1.8037

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	58.48326	29.24163	2.47	0.0896
Error	107	1267.988	11.85035		
Corrected Total	109	1326.471			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	63.48252	1.62036	18189	1534.91	<.0001
MA	0.14924	0.0672	58.44663	4.93	0.0285
ART	-0.06025	0.05754	12.99449	1.1	0.2974

Bounds on condition number: 1.3216, 5.2864

Forward Selection: Step 3

Variable RCT Entered: R-Square = 0.0512 and C(p) = 3.0147

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	67.93922	22.64641	1.91	0.1328
Error	106	1258.532	11.87294		
Corrected Total	109	1326.471			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F

Intercept	62.90059	1.74808	15373	1294.75	<.0001
MA	0.12201	0.07386	32.39391	2.73	0.1015
RCT	0.08706	0.09756	9.45596	0.8	0.3742
ART	-0.06231	0.05764	13.8766	1.17	0.2821

Bounds on condition number: 1.5936, 12.646

No other variable met the 0.5000 significance level for entry into the model.

Summary of Forward Selection

Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	MA	1	0.0343	0.0343	0.888	3.84	0.0528
2	ART	2	0.0098	0.0441	1.8037	1.1	0.2974
3	RCT	3	0.0071	0.0512	3.0147	0.8	0.3742

Race=1

Number of Observations R	110
Number of Observations U	110

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	67.93922	22.64641	1.91	0.1328
Error	106	1258.532	11.87294		
Corrected Total	109	1326.471			

Root MSE	3.44571	R-Square	0.0512
Dependent Mean	66.39091	Adj R-Sq	0.0244
Coeff Var	5.19004		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	62.90059	1.74808	35.98	<.0001	0
MA	1	0.12201	0.07386	1.65	0.1015	0.19728
RCT	1	0.08706	0.09756	0.89	0.3742	0.09619
ART	1	-0.06231	0.05764	-1.08	0.2821	-0.11768

Correlation of Estimates

Variable	Intercept	MA	RCT	ART
Intercept	1	-0.5166	-0.373	-0.0844
MA	-0.5166	1	-0.4132	-0.4323
RCT	-0.373	-0.4132	1	-0.0401
ART	-0.0844	-0.4323	-0.0401	1

Collinearity Diagnostics

Number	Eigenvalue	Condition Index	Proportion of Variation			
			Intercept	MA	RCT	ART
1	3.88021	1	0.00229	0.0017	0.00285	0.00528
2	0.07157	7.36287	0.04737	0.00401	0.10877	0.83923
3	0.02934	11.49943	0.48762	0.03646	0.81574	0.0329
4	0.01887	14.33821	0.46272	0.95783	0.07264	0.1226

Test of First and Second Moment Specification		
DF	Chi-Square	Pr > ChiSq
9	7.72	0.563

Durbin-Watson D	2.001
Number of Observations	110
1st Order Autocorrelation	-0.009

Race=2

Number of Observations R	182
Number of Observations U	182

Forward Selection: Step 1

Variable MA Entered: R-Square = 0.0512 and C(p) = 1.5914

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	138.2989	138.2989	9.72	0.0021
Error	180	2561.773	14.23207		
Corrected Total	181	2700.072			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	60.28275	1.08001	44341	3115.54	<.0001
MA	0.14282	0.04582	138.2989	9.72	0.0021

Bounds on condition number: 1, 1

Forward Selection: Step 2

Variable RCT Entered: R-Square = 0.0586 and C(p) = 2.1948

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	158.2211	79.11057	5.57	0.0045
Error	179	2541.851	14.20028		
Corrected Total	181	2700.072			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	59.57331	1.23392	33100	2330.94	<.0001
MA	0.11334	0.05209	67.22511	4.73	0.0309
RCT	0.09414	0.07948	19.92223	1.4	0.2378

Bounds on condition number: 1.2957, 5.1828

Forward Selection: Step 3

Variable ART Entered: R-Square = 0.0649 and C(p) = 3.0115

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	175.0997	58.36658	4.11	0.0075
Error	178	2524.972	14.18524		
Corrected Total	181	2700.072			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	59.4808	1.23618	32842	2315.23	<.0001
MA	0.07444	0.06311	19.74018	1.39	0.2397
RCT	0.09915	0.07957	22.02488	1.55	0.2144
ART	0.05819	0.05334	16.87859	1.19	0.2768

Bounds on condition number: 1.9035, 14.3

No other variable met the 0.5000 significance level for entry into the model.

Summary of Forward Selection							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	MA	1	0.0512	0.0512	1.5914	9.72	0.0021
2	RCT	2	0.0074	0.0586	2.1948	1.4	0.2378
3	ART	3	0.0063	0.0649	3.0115	1.19	0.2768

Race=2

Number of Observations R	182
Number of Observations U	182

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	175.0997	58.36658	4.11	0.0075
Error	178	2524.972	14.18524		
Corrected Total	181	2700.072			
Root MSE	3.76633	R-Square	0.0649		
Dependent Mean	63.53462	Adj R-Sq	0.0491		
Coeff Var	5.928				

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	59.4808	1.23618	48.12	<.0001	0
MA	1	0.07444	0.06311	1.18	0.2397	0.11797
RCT	1	0.09915	0.07957	1.25	0.2144	0.10298
ART	1	0.05819	0.05334	1.09	0.2768	0.09885

Correlation of Estimates				
Variable	Intercept	MA	RCT	ART
Intercept	1	-0.381	-0.4874	-0.0686
MA	-0.381	1	-0.4261	-0.5651

RCT	-0.4874	-0.4261	1	0.0577
ART	-0.0686	-0.5651	0.0577	1

Collinearity Diagnostics

Number	Eigenvalue	Condition Index	Proportion of Variation			
			Intercept	MA	RCT	ART
1	3.84101	1	0.00334	0.00234	0.00347	0.00608
2	0.09774	6.26872	0.06688	4.88E-05	0.11308	0.65407
3	0.035	10.47567	0.82622	0.00455	0.65049	0.00311
4	0.02625	12.09688	0.10357	0.99306	0.23295	0.33673

Dependent Variable: Average Core Courses

Race=2

Test of First and Second Moment Specification		
DF	Chi-Square	Pr > ChiSq
9	10.93	0.2802
Durbin-Watson D	2.04	
Number of Observations	182	
1st Order Autocorrelation	-0.027	

Race=3

Number of Observations R	449
Number of Observations U	449

Forward Selection: Step 1

Variable RCT Entered: R-Square = 0.0593 and C(p) = 12.1729

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	440.64	440.64	28.2	<.0001
Error	447	6984.601	15.62551		
Corrected Total	448	7425.241			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	62.89013	0.80491	95391	6104.82	<.0001
RCT	0.25086	0.04724	440.64	28.2	<.0001

Bounds on condition number: 1, 1

Forward Selection: Step 2

Variable CRT Entered: R-Square = 0.0822 and C(p) = 3.0458

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	610.6376	305.3188	19.98	<.0001

Error	446	6814.603	15.27938		
Corrected Total	448	7425.241			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	60.87231	0.99974	56646	3707.36	<.0001
RCT	0.18974	0.05018	218.4658	14.3	0.0002
CRT	0.1182	0.03544	169.9976	11.13	0.0009

Bounds on condition number: 1.1539, 4.6155

Forward Selection: Step 3

Variable ART Entered: R-Square = 0.0853 and C(p) = 3.5590

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	633.3533	211.1178	13.83	<.0001
Error	445	6791.888	15.26267		
Corrected Total	448	7425.241			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	60.93503	1.00052	56613	3709.25	<.0001
RCT	0.18042	0.05073	193.035	12.65	0.0004
ART	0.04412	0.03616	22.71571	1.49	0.2231
CRT	0.08969	0.04244	68.17644	4.47	0.0351

Bounds on condition number: 1.6564, 13.354

Forward Selection: Step 4

Variable MA Entered: R-Square = 0.0864 and C(p) = 5.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	641.893	160.4733	10.5	<.0001
Error	444	6783.348	15.27781		
Corrected Total	448	7425.241			
Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	60.80679	1.0156	54767	3584.74	<.0001
MA	0.03469	0.04639	8.53969	0.56	0.4551
RCT	0.16387	0.05537	133.8065	8.76	0.0032
ART	0.03267	0.03929	10.56625	0.69	0.4061
CRT	0.07614	0.04616	41.56893	2.72	0.0998

Bounds on condition number: 2.5319, 31.193

All variables have been entered into the model.

Summary of Forward Selection							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	RCT	1	0.0593	0.0593	12.1729	28.2	<.0001
2	CRT	2	0.0229	0.0822	3.0458	11.13	0.0009
3	ART	3	0.0031	0.0853	3.559	1.49	0.2231
4	MA	4	0.0012	0.0864	5	0.56	0.4551

Race=3

Number of Observations R	449
Number of Observations U	449

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	641.893	160.4733	10.5	<.0001
Error	444	6783.348	15.27781		
Corrected Total	448	7425.241			
Root MSE	3.90868	R-Square	0.0864		
Dependent Mean	67.04811	Adj R-Sq	0.0782		
Coeff Var	5.82967				

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	60.80679	1.0156	59.87	<.0001	0
MA	1	0.03469	0.04639	0.75	0.4551	0.05396
RCT	1	0.16387	0.05537	2.96	0.0032	0.15913
ART	1	0.03267	0.03929	0.83	0.4061	0.05204
CRT	1	0.07614	0.04616	1.65	0.0998	0.1047

Correlation of Estimates					
Variable	Intercept	MA	RCT	ART	CRT
Intercept	1	-0.1689	-0.3855	0.1125	-0.4166
MA	-0.1689	1	-0.3997	-0.3897	-0.3925
RCT	-0.3855	-0.3997	1	0.0286	-0.0272
ART	0.1125	-0.3897	0.0286	1	-0.3136
CRT	-0.4166	-0.3925	-0.0272	-0.3136	1

Collinearity Diagnostics							
Number	Eigenvalue	Condition Index	Proportion of Variation				
			Intercept	MA	RCT	ART	CRT
1	4.87035	1	0.00135	0.000811	0.00156	0.00228	0.000962
2	0.06501	8.65545	0.08735	0.000165	0.10846	0.55842	0.000249
3	0.03029	12.68082	0.26765	0.00218	0.67683	0.0412	0.14739
4	0.02013	15.55467	0.58466	0.21731	0.00467	0.37861	0.32418
5	0.01422	18.50848	0.05899	0.77954	0.20849	0.01949	0.52722

Race=3

Test of First and Second

Moment Specification		
DF	Chi-Square	Pr > ChiSq
14	10.73	0.7067
Durbin-Watson D	1.962	
Number of Observations	449	
1st Order Autocorrelation	0.018	

Appendix G: Logistic Regression Results

The LOGISTIC Procedure					
Model Information					
Data Set	WORK.SORTTEMPTABLESORTED				
Response Variable	Success/Not				
Number of Response Levels	2				
Model	binary logit				
Optimization Technique	Fisher's scoring				
Number of Observations Read	741				
Number of Observations Used	741				
Response Profile					
Ordered	**Success/Not**	**Total**			
Value		**Frequency**			
1	0	312			
2	1	429			
Probability modeled is Success/Not='1'					
.					
Forward Selection Procedure					
Step 0. Intercept entered:					
Model Convergence Status					
nvergence criterion (GCONV=1E-8) satisfied.					
-2 Log L	=	1008.693			
Residual Chi-Square Test					
Chi-Square	**DF**	**Pr > ChiSq**			
2.9893	4	0.5596			
Note: No (additional) effects met the 0.05 significance level for entry into the model.					
Analysis of Maximum Likelihood Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	0.3185	0.0744	18.3184	<.0001
Profile Likelihood Confidence Interval for Parameters					
Parameter	**Estimate**	**95% Confidence Limits**			
Intercept	0.3185	0.1731	0.4649		

Appendix H: 2 Independent Sample T-test

The TTEST Procedure

Variable: MA

Prefered Language	N	Mean	Std Dev	Std Err	Minimum	Maximum
1	492	26.9797	6.4432	0.2905	4	41
2	249	25.3333	6.5591	0.4157	6	39
Diff (1-2)		1.6463	6.4823	0.5041		

Prefered Language	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		26.9797	26.4089	27.5504	6.4432	6.0642	6.8731
2		25.3333	24.5147	26.152	6.5591	6.0291	7.192
Diff (1-2)	Pooled	1.6463	0.6566	2.6361	6.4823	6.168	6.8306
Diff (1-2)	Satterthwa	1.6463	0.65	2.6427			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	739	3.27	0.0011
Satterthwa	Unequal	490.32	3.25	0.0012

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	248	491	1.04	0.7368

Variable: RCT

Prefered Language	N	Mean	Std Dev	Std Err	Minimum	Maximum
1	492	16.2175	4.1865	0.1887	0	26
2	249	15.4378	3.6584	0.2318	5	24
Diff (1-2)		0.7797	4.0171	0.3124		

Prefered Language	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		16.2175	15.8466	16.5883	4.1865	3.9403	4.4659
2		15.4378	14.9811	15.8944	3.6584	3.3629	4.0114
Diff (1-2)	Pooled	0.7797	0.1664	1.3931	4.0171	3.8223	4.2329
Diff (1-2)	Satterthwa	0.7797	0.1925	1.3669			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	739	2.5	0.0128
Satterthwa	Unequal	561.16	2.61	0.0093

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	491	248	1.31	0.0166

Variable: ART

Prefered Language	N	Mean	Std Dev	Std Err	Minimum	Maximum
1	492	18.4654	6.6526	0.2999	0	30
2	249	16.7189	6.4871	0.4111	0	30
Diff (1-2)		1.7466	6.5975	0.5131		

Prefered Language	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		18.4654	17.8762	19.0547	6.6526	6.2613	7.0965
2		16.7189	15.9092	17.5286	6.4871	5.9629	7.113
Diff (1-2)	Pooled	1.7466	0.7393	2.7539	6.5975	6.2777	6.952
Diff (1-2)	Satterthwa	1.7466	0.7468	2.7463			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	739	3.4	0.0007
Satterthwa	Unequal	509.37	3.43	0.0006

Equality of Variances				

Method	Num DF	Den DF	F Value	Pr > F
Folded F	491	248	1.05	0.6572

Variable: CRT

Prefered Language	N	Mean	Std Dev	Std Err	Minimum	Maximum
1	492	24.8841	5.8617	0.2643	5	35
2	249	23.6225	5.7797	0.3663	4	35
Diff (1-2)		1.2617	5.8343	0.4537		

Prefered Language	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		24.8841	24.3649	25.4034	5.8617	5.5169	6.2528
2		23.6225	22.9011	24.3439	5.7797	5.3128	6.3374
Diff (1-2)	Pooled	1.2617	0.3709	2.1524	5.8343	5.5514	6.1478
Diff (1-2)	Satterthwa	1.2617	0.3743	2.149			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	739	2.78	0.0056
Satterthwa	Unequal	504.36	2.79	0.0054

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	491	248	1.03	0.8081

Variable: Average Core Courses

Prefered Language	N	Mean	Std Dev	Std Err	Minimum	Maximum
1	492	66.628	4.0866	0.1842	55.3	79
2	249	65.0197	4.2287	0.268	56	76.9
Diff (1-2)		1.6084	4.1349	0.3216		

Prefered Language	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		66.628	66.2661	66.99	4.0866	3.8463	4.3593
2		65.0197	64.4919	65.5475	4.2287	3.887	4.6367
Diff (1-2)	Pooled	1.6084	0.9771	2.2397	4.1349	3.9344	4.357
Diff (1-2)	Satterthwa	1.6084	0.9694	2.2474			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	739	5	<.0001
Satterthwa	Unequal	483.31	4.95	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	248	491	1.07	0.526

Appendix I: One-Way Analysis of Variance

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Race	3	1 2 3
Number of	741	
Number of	741	

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	3282.14173	1641.071	42.92	<.0001
Error	738	28219.0998	38.23726		
Corrected	740	31501.2416			

R-Square	Coeff Var	Root MSE	MA Mean
0.104191	23.39939	6.183629	26.42645

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Race	2	3282.14173	1641.071	42.92	<.0001

Dependent Variable: RCT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	492.17032	246.0852	15.75	<.0001
Error	738	11533.3601	15.62786		
Corrected	740	12025.5304			

R-Square	Coeff Var	Root MSE	RCT Mean
0.040927	24.77652	3.953209	15.95547

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Race	2	492.170317	246.0852	15.75	<.0001

Dependent Variable: ART

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1309.46713	654.7336	15.41	<.0001
Error	738	31361.6017	42.4954		
Corrected	740	32671.0688			

R-Square	Coeff Var	Root MSE	ART Mean
0.04008	36.46186	6.518849	17.87854

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Race	2	1309.46713	654.7336	15.41	<.0001

Dependent Variable: CRT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	2063.35014	1031.675	32.6	<.0001
Error	738	23354.7254	31.64597		
Corrected	740	25418.0756			

R-Square	Coeff Var	Root MSE	CRT Mean
0.081176	22.99849	5.625475	24.46019

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Race	2	2063.35014	1031.675	32.6	<.0001

Dependent Variable: Average Core Courses

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1610.58205	805.291	51.9	<.0001
Error	738	11451.7837	15.51732		
Corrected	740	13062.3658			

R-Square	Coeff Var	Root MSE	Average Core Courses Mean
0.123299	5.96058	3.939203	66.08758

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Race	2	1610.58205	805.291	51.9	<.0001

The ANOVA Procedure

Levene's Test for Homogeneity of MA Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Race	2	6544.8	3272.4	0.65	0.523
Error	738	3722317	5043.8		

Levene's Test for Homogeneity of RCT Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Race	2	114.1	57.0461	0.1	0.9011
Error	738	404306	547.8		

Levene's Test for Homogeneity of ART Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Race	2	157.7	78.8272	0.02	0.9759
Error	738	2385078	3231.8		

Levene's Test for Homogeneity of CRT Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Race	2	106.2	53.1247	0.02	0.9796
Error	738	1905976	2582.6		

Levene's Test for Homogeneity of Average Core Courses Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Race	2	1864.5	932.3	2.16	0.1166
Error	738	319160	432.5		