School results and access test results as indicators of first-year performance at university

A Bothma*   HL Botha†   NJ Le Roux‡

Abstract

The goals set by the National Plan for Higher Education, the fact that many schools are still severely disadvantaged as well as far-reaching changes in the school system demand that South African universities urgently reconsider their admission procedures. Redesigning admission procedures calls for a thorough understanding of the interrelationships between school marks, results in existing access tests and first-year university performance. These interrelationships were statistically investigated in the case of the 1999, 2000 and 2001 intake groups, who were compelled to write access tests before being admitted to Stellenbosch University. The results of this investigation confirm an alarming degree of unpreparedness among many prospective students regarding what is expected of them at university. This is aggravated by school marks creating a totally unrealistic expectation of performance in the first year at university. It is emphasised that schools and authorities dealing with admission of prospective students at universities should be cognisant of the findings reported here. Furthermore, the statistical analyses demonstrate several novel techniques for investigating the interrelationship between school marks, access test results and university performance.

1   Introduction

The National Plan for Higher Education (NPHE) [4] outlines the framework and mechanisms for implementing and realising the policy goals of the Education White Paper 3: A Programme for the Transformation of Higher Education [1]. The key challenges facing the South African higher education (HE) system are outlined in the White Paper as follows: "to redress past inequalities and to transform the higher education system to serve social order, to meet pressing national needs, and to respond to new realities and opportunities" [1].

---

*Department of Statistics and Actuarial Science, University of Stellenbosch, Private Bag X1, Matieland, 7602, Stellenbosch, South Africa

†Academic Support, University of Stellenbosch, Private Bag X1, Matieland, 7602, Stellenbosch, South Africa

‡Corresponding author: Department of Statistics and Actuarial Science, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa, email: `njlr@sun.ac.za`

The following two of the five policy goals of the NPHE are highlighted:

- Producing the graduates needed for social and economic development in South Africa; and

- Achieving equity in the South African HE system.

There is a great shortage of high-level professional and managerial skills in South Africa. This, along with other labour market trends (listed in the NPHE), clearly stresses the need for the HE system to produce more graduates. The HE system is at present producing insufficient numbers of graduates. The average graduation rate remained at 15% between 1993 and 1998. Therefore, the following priority was identified in order to achieve the first goal of the NPHE: to increase the participation rate and number of graduates in HE in order to meet the demand for high-level skills through a balanced production of graduates in different fields of study. In order to meet their second goal the NPHE stipulates that the student profiles should progressively reflect the demographic realities of South African society.

These goals set by the NPHE, the fact that many schools are still severely disadvantaged and changes in the school system make it necessary for universities to reconsider their admission procedures as a matter of urgency. The changes in the school system especially include the introduction of the Further Education and Training Certificate (FETC) and the expected discontinuation of the matriculation endorsement. New procedures are required to determine the suitability of applicants who do not meet the minimum criteria for automatic admission (currently matriculation endorsement) as well as the minimum criteria for automatic admission into different degree programmes. In the light of the above, it seems that there is an urgent need for measures and admission criteria that can predict success at university to some extent. Furthermore, to promote access equity these measures and procedures should be able to identify students who have the ability (potential) to be successful at university level.

Previous studies in South Africa on school results have shown that Grade 12 means correlate better with first-year performance than any other psychometric predictor [6, 7, 8]. However, the experience of universities is that students with good Grade 12 marks do not necessarily pass their first year.

In order to adapt to these changing circumstances Stellenbosch University (SU) designed a battery of tests (access tests) to extend its admission procedures. This battery of tests consists of academic proficiency tests in Afrikaans and English, as well as a Mathematics, a Science, a Numeracy and a Thinking Skills test. The tests are aimed at Grade 12 learners, and although the main focus is on existing knowledge and skills necessary to study successfully at a university, elements of the battery provide some measure of potential (especially the Thinking Skills test). Each learner writes three of the six tests based on chosen field of study. The main goal of the access tests is to determine whether applicants are adequately prepared for university studies and to allow certain candidates who are not in possession of matriculation endorsement, but who show potential to be successful at university, to enrol at SU. These tests are therefore used to see whether a prospective student's level of academic preparation is sufficient for study at SU, and to advise and

counsel students on the basis of the test results. Student performance in the above access tests is not only of utmost importance to SU, but indeed for all concerned about whether schools are producing sufficient numbers of adequately prepared candidates for study at university.

In view of the above it is apparent that the ability of these access tests to predict success at university should be investigated. The main focus of this investigation is to examine

- to what extent access tests provide information regarding the preparedness of students for their intended field of study;

- the relationships between access tests and school results;

- the ability of access tests to discriminate between potentially successful and unsuccessful students;

- the relationship between access test results and first year performance;

- the use of new techniques to examine the distributions of the school results, access tests and first-year performance variables.

In order to address these aims data sets were compiled of all prospective students who were required to write access tests and eventually enrolled at SU in 1999, 2000 and 2001, respectively. The data sets that were subjected to statistical analyses consisted of school results, access test results and first-year performance at university.

The results of the above statistical analyses are reported in two related articles. The remainder of this article consists of three sections. Firstly, the data sets analysed are examined more closely. This is followed by a section in which access test results, school results and first-year university performance are described and analysed in terms of boxplots and correlations. A final discussion of the results obtained by the boxplot and correlation analyses together with some recommendations concludes the article. In the second article [3] we take a closer look at the differences and similarities found among the statistical distributions of the access test variables, school result variables and first-year performance, respectively. More sophisticated density estimates are considered in this second article and it is demonstrated how these estimates provide detailed descriptions of the underlying characteristics of the above variables.

2   Data sets

The first group of students required to write access tests at SU, enrolled in 1999. This group and the subsequent 2000 and 2001 groups will be referred to as the **1999 intake group**, the **2000 intake group** and the **2001 intake group** respectively. The 1999 intake group consists of all prospective students with a school Grade 11 final average mark of less than 60% or Grade 12 final average examination mark of less than 57%. Stricter requirements were applied for the 2000 intake group. Prospective students with a Grade 11 final average mark of less than 70%, or a Grade 12 final average examination mark of less than 68.57%, were then compelled to write the access tests. For the 2001

intake all prospective Health Sciences students with a Grade 11 final average mark, or Grade 12 final average examination mark of less than 85%, had to write the access tests. Prospective students in all other faculties had to write access tests if their Grade 11 or Grade 12 final average marks were less than 70% or 68.57%, respectively. Separate data sets for each of the 1999, 2000 and 2001 intake groups were compiled. Each of the data sets consists of access test variables, school result variables and a first year performance variable.

The access test battery consists of the following individual tests:

1. Academic Language Proficiency in Afrikaans (Afr),

2. Academic Language Proficiency in English (Eng),

3. Mathematics (Maths),

4. Physical Science (Science),

5. Numeracy Skills (Numer) and

6. Thinking Skills (Think).

A prospective student has to write a selection of three tests according to the faculty in which he/she plans to study. Table 1 summarises these different test batteries, each consisting of three individual tests.

Test Battery	Faculty	Individual Access Tests
1A 1E	Health Sciences, Natural Sciences, Agricultural and Forestry Sciences, Engineering	Maths, Science and Afr Maths, Science and Eng
2A 2E	Economic and Management Sciences	Maths, Numer and Afr Maths, Numer and Eng
3	Arts, Theology, Law and Education	Think, Afr and Eng

Table 1: *Composition of access test batteries.*

After the access test results of the 1999 intake group were analysed, it became evident that the Thinking Skills test was too easy and the Mathematics, Physical Science and Numeracy Skills tests were too difficult. These tests were adjusted accordingly in 1999, before being administered to the 2000 intake group. This caused a drop in the Thinking Skills average from 1999 to that obtained in 2000, and a corresponding increase in the averages of Mathematics, Physical Science and Numeracy Skills. The effects of these changes to the access tests will be referred to as the **access test adjustment effects** in this article.

Another phenomenon that will often be referred to in the data analysis is the change in the construction of the 2001 intake group. The 2001 intake group includes a greater number of students with higher Grade 11 and Grade 12 final average marks, which will be referred to as the **effect of the Health Sciences 2001 students**.

The following school results variables are included in the data sets: the Grade 12 Mathematics mark (Maths.12), the Grade 12 Afrikaans mark (Afr.12), the Grade 12 English mark (Eng.12), Grade 11 average mark (Ave.11) and the Grade 12 average mark (Ave.12). The variable Ave.11 refers to the final average mark a student received at the end of Grade 11. This is often used as an initial average mark to apply for admission to the University. The variable Ave.12, on the other hand, refers to the official matriculation average mark a student receives at the end of his/her school career.

The last variable included in each of the data sets is the first-year weighted university mark (FYWUM) that was computed for each student. This mark is calculated as follows: First the mark achieved for each semester or year module is multiplied by the corresponding credit value for that module. The aggregate of these products divided by the aggregate of the credit values of all *required* modules taken comprises the FYWUM.

It is apparent from Table 1 that each student wrote only one of five possible access test combinations. The data of each intake group were divided into the five data sets: Test Battery 1A, Test Battery 1E, Test Battery 2A, Test Battery 2E and Test Battery 3 consisting of the corresponding access test variables (*cf.* Table 1), the five school result variables and FYWUM. Due to the small number of students in some of the above data sets, the possibility of merging the Afrikaans and English test batteries for each of the three years was investigated. A comparison of corresponding Afrikaans and English data sets revealed only minor disparities, resulting in the merging of the Test Battery 1A and Test Battery 1E data sets, as well as the Test Battery 2A and Test Battery 2E data sets. The new Test Battery 1 and Test Battery 2 data sets each contains a variable Language (Lang) which replaces the original variables Afr and Eng. The variable AT.Ave, representing the average of the three access tests marks, was added to each of the data sets.

Any student with an incomplete record regarding the above variables was excluded from the statistical analyses. Table 2 shows the number of records in each of the three test batteries for the different intake groups, together with the respective variables. In order to uniquely define each record an identification number (Studnum) was allocated to each record.

Test Battery	Number of records			Variables
	1999	2000	2001	
TB 1	81	184	511	Studnum, Maths, Science, Lang, AT.Ave, Maths.12, Afr.12, Eng.12, Ave.11, Ave.12, FYWUM
TB 2	81	260	230	Studnum, Maths, Numer, Lang, AT.Ave, Maths.12, Afr.12, Eng.12, Ave.11, Ave.12, FYWUM
TB 3	78	208	246	Studnum, Think, Afr, Eng, AT.Ave, Afr.12, Eng.12, Ave.11, Ave.12, FYWUM

Table 2: *Number of records and variables associated with the respective data sets.*

3   Data analysis

In this section exploratory techniques are used to obtain a better understanding of the shape and characteristics of the statistical distribution of each of the variables described in the previous section and to detect the presence of outliers. Boxplots are constructed to provide univariate graphical displays of the variables. In addition, statistical hypothesis tests to compare various groups of variables univariately are discussed.

Since similar patterns of statistical results often occur for the three intake groups of a Test Battery, only results of representative intake groups are reported here.

3.1   Boxplots of the variables included in the data analysis

Notched boxplots showing outliers, maximum and minimum values (excluding outliers), together with the quartiles were constructed. An outlier is defined here as a point beyond a standard span of the quartiles, where a standard span is equal to 1.5 times the interquartile range of the data set. The whiskers are drawn to the nearest value not beyond a standard span from the quartiles.

Notched boxplots are characterised by notches that surround the medians, demarcating an approximate 95% confidence interval. The notches provide a measure of an approximate test of the significance of differences between two or more medians. Specifically, if the notches of two medians do not overlap, the medians are considered significantly different at an approximate 5% significance level [5]. If the medians of two or more groups are compared simultaneously, the overall significance level is unknown. The Bonferroni inequality [2] may be used to compute an upper bound for the associated overall significance level, *i.e.* a lower bound for the 95% confidence interval.

3.1.1   Test battery 1

Figure 1 displays the univariate boxplots of each of the continuous variables in the Test Battery 1 data set for the 2000 intake group. The means and standard deviations of the Test Battery 1 data sets for 1999, 2000 and 2001 are summarised in Tables 3 and 4, respectively.

Year	Maths	Science	Lang	AT. Ave	Maths. 12	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
1999	35.29	37.39	61.20	44.63	59.02	69.31	64.31	68.92	64.66	39.39
2000	40.09	50.62	67.08	52.60	59.21	68.93	64.87	68.82	66.78	34.89
2001	43.32	52.64	70.69	55.56	66.25	73.89	69.13	77.45	75.77	50.03

Table 3: *Means of the variables of the Test Battery 1 data set for the 1999, 2000 and 2001 intake groups.*

It follows from the boxplots in Figure 1 that the medians of Maths and FYWUM are less than 40%, while the medians of Lang, Afr.12, Eng.12, Ave.11 and Ave.12 are in excess of 65%. From Table 3 the similarity in the medians and means of these variables is apparent for the 2000 intake group, and the same holds for the 1999 and 2001 intake groups.

Year	Maths	Science	Lang	AT. Ave	Maths. 12	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
1999	15.24	12.45	13.38	9.86	12.13	9.52	7.71	9.23	7.46	17.48
2000	14.40	14.23	11.68	10.41	13.17	8.53	8.39	8.08	6.83	17.14
2001	18.49	16.31	12.65	12.97	14.30	8.92	8.76	11.82	10.83	22.02

Table 4: *Standard deviations of the variables of the Test Battery 1 data set for the 1999, 2000 and 2001 intake groups.*

A general tendency is evident in Table 3: the means of Maths and Science correspond reasonably closely with those of FYWUM, but the means of the school results variables are substantially higher than those of FYWUM.
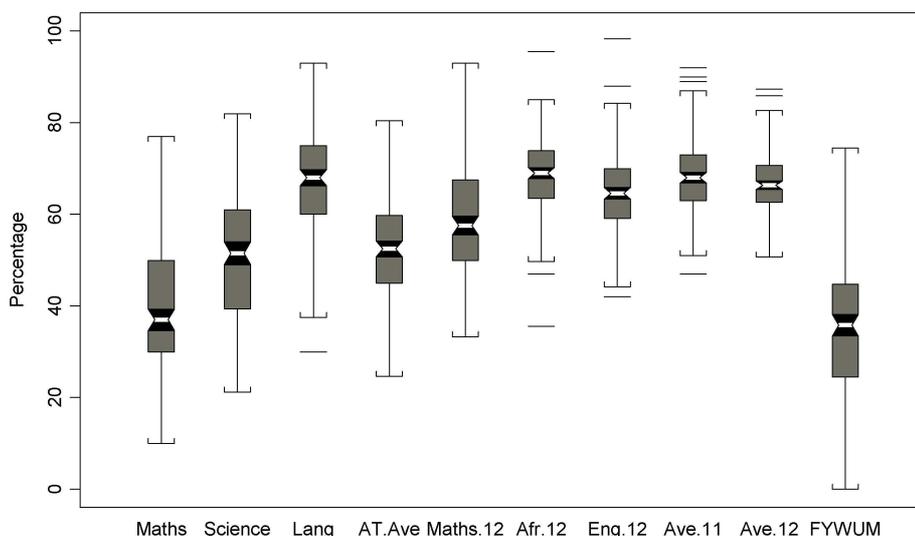
Figure 1: *Notched boxplots of the variables of the Test Battery 1 data set for the 2000 intake group.*

Since it can be misleading to investigate the means without providing some measure of variability, the corresponding standard deviations are given in Table 4. It is clear that access test variables show greater standard deviations than the school result variables. Only Grade 12 Mathematics (Maths.12) show variation that is comparable to that of the access test variables. Accordingly, the variation in Maths exceeds that of any of the other access test variables. The sizes of the boxes in Figure 1, confirm these trends in the variations. The striking differences between Ave.12 and FYWUM are a matter of concern. The question arises as to why a group of students, obtaining approximately similar school marks (students with Grade 12 marks mostly between 60% and 70%), obtain university marks that are, on average, considerably lower and have far more variation than their school marks. In 2001 the standard deviations of Maths, Science, Ave.11, Ave.12 and FYWUM show a substantial increase. The wider spectrum of prospective Health Sciences students of 2001 has an inflating effect on the variation of several variables.

From the boxplots of Figure 1 it is clear that the confidence interval of the median of FYWUM overlaps only with the confidence interval of the median of access tests variable Maths. This indicates that the location of FYWUM differs from the location of each of the school result variables at approximately a 5% significance level. Therefore, the access test variables give a better idea of the location of FYWUM than any of the school result variables. Indeed, the difference between the means of Ave.12 and FYWUM is at least 25 percentage points for all three intake groups. Moreover, the third quartile of FYWUM is lower than or equal to the first quartile of Ave.11, Ave.12, Afr.12 and Eng.12. These findings show that prospective students could find it extremely difficult maintaining their school performance at university. The results displayed in Figure 1 and Table 3 are not isolated cases — these are the general tendencies found in all data sets considered in this article.

Table 3 shows an increase in the means of all variables in 2001, in particular FYWUM. This increase may be attributed to the *effect of the Health Sciences 2001 students*. This makes the Test Battery 1 data set of 2001 of special importance, since it can be used to analyse the effect of the wider spectrum of students on the relationship between the access tests and FYWUM, as well as the relationship between the access tests and school results. Moreover, the means of Science and Maths improved from 1999 to 2000 due to the *access test adjustment effect*.

3.1.2   Test battery 2

Notched boxplots are displayed in Figure 2 to compare the relevant variables of the Test Battery 2 data set for the 2000 intake group. Furthermore, Tables 5 and 6 contain the means and standard deviations of these variables, respectively.

Year	Maths	Numer	Lang	AT. Ave	Maths. 12	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
1999	33.69	38.72	61.02	44.48	58.36	63.02	62.36	63.42	62.08	31.85
2000	36.54	44.60	63.80	48.39	61.11	68.13	64.16	66.06	65.83	34.45
2001	31.04	42.73	63.25	45.68	64.66	68.96	63.99	66.30	66.75	32.83

Table 5: *Means of the variables of the Test Battery 2 data set for the 1999, 2000 and 2001 intake groups.*

Year	Maths	Numer	Lang	AT. Ave	Maths. 12	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
1999	16.33	14.82	13.23	10.82	16.22	12.26	9.95	7.75	9.87	18.97
2000	13.16	16.02	13.37	10.63	13.13	8.66	6.81	7.39	5.79	15.72
2001	16.07	16.27	12.51	11.94	12.70	9.17	8.07	7.87	7.25	15.13

Table 6: *Standard deviations of the variables of the Test Battery 2 data set for the 1999, 2000 and 2001 intake groups.*

Figure 2 indicates that the medians of Maths and FYWUM are less than 38%, while the medians of Lang, Afr.12, Eng.12, Ave.11 and Ave.12 are in excess of 62%. Table

5 demonstrates that the means of Maths, Numer and FYWUM correspond relatively closely for the three intake groups, while the means of the school result variables are again considerably higher than those of FYWUM. Table 6 indicates that the school result variables generally show less variation than the access test variables and FYWUM, but the standard deviation of Maths.12 is comparable to those of the access test variables. It is striking, once again, that despite very similar school marks (as illustrated in Figure 2), the corresponding first-year marks are not only substantially lower, but also have noticeably more variation.
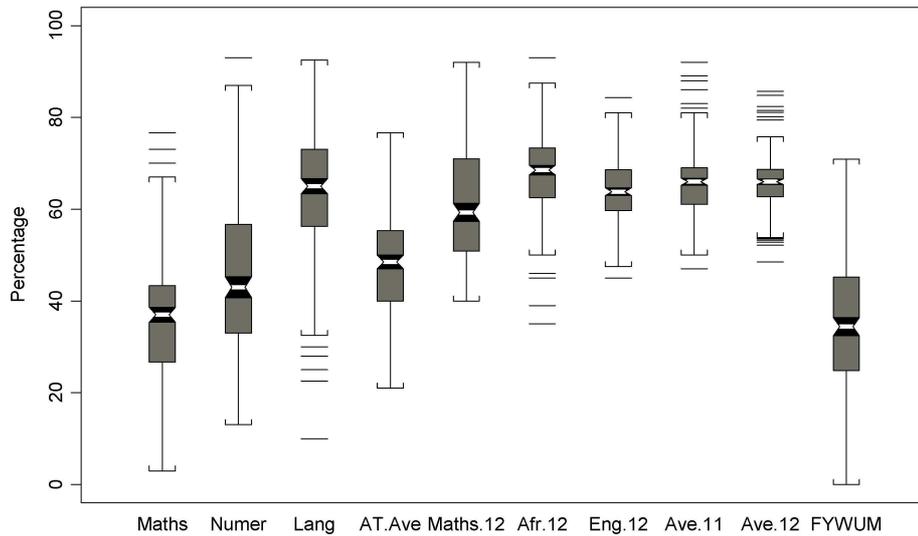
Figure 2: *Notched boxplots of the variables of the Test Battery 2 data set for the 2000 intake groups.*

Figure 2 demonstrates that the confidence interval of the median of FYWUM overlaps only with the confidence interval of Maths, implying that the medians of FYWUM and each of the school result variables differ at approximately a 5% significance level. Therefore, some access test results give a good indication of the location of FYWUM, while Grade 12 and Grade 11 final marks create unrealistic expectations of university performance. This is emphasised by Figure 2, which indicates that the third quartile of FYWUM is lower than the first quartile of Ave.11, Ave.12, Afr.12, Eng.12 and Maths.12. The increase in the medians of Maths and Numer from 1999 to 2000 may be explained by the *access test adjustment effect*.

3.1.3 Test battery 3

In Figure 3 the notched boxplots of the Test Battery 3 data set for the 2000 intake group are displayed. The means and standard deviations of these variables are listed in Tables 7 and 8, respectively.

It is noted for Test Battery 3 that the means (and medians) of the access test variables are somewhat higher than that of FYWUM, with Think on average ten percentage points

Year	Think	Afr	Eng	AT. Ave	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
1999	61.12	59.10	53.27	57.83	67.86	65.18	61.23	62.03	37.67
2000	50.06	64.20	54.92	56.23	70.31	65.17	64.28	63.71	40.37
2001	51.81	59.78	54.31	55.26	70.47	66.14	64.80	64.87	40.21

Table 7: *Means of the variables of the Test Battery 3 data set for the 1999, 2000 and 2001 intake groups.*

Year	Think	Afr	Eng	AT. Ave	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
1999	16.13	16.43	13.62	10.21	9.57	6.68	6.94	7.69	14.73
2000	16.16	13.68	14.63	11.39	7.58	7.03	7.94	5.91	18.11
2001	16.28	15.61	13.71	10.86	7.81	7.54	7.31	6.86	17.41

Table 8: *Standard deviations of the variables of the Test Battery 3 data set for the 1999, 2000 and 2001 intake groups.*

higher than FYWUM. Table 8 reveals that the standard deviations of the school result variables are once again appreciably lower than those of the three access test variables and FYWUM.
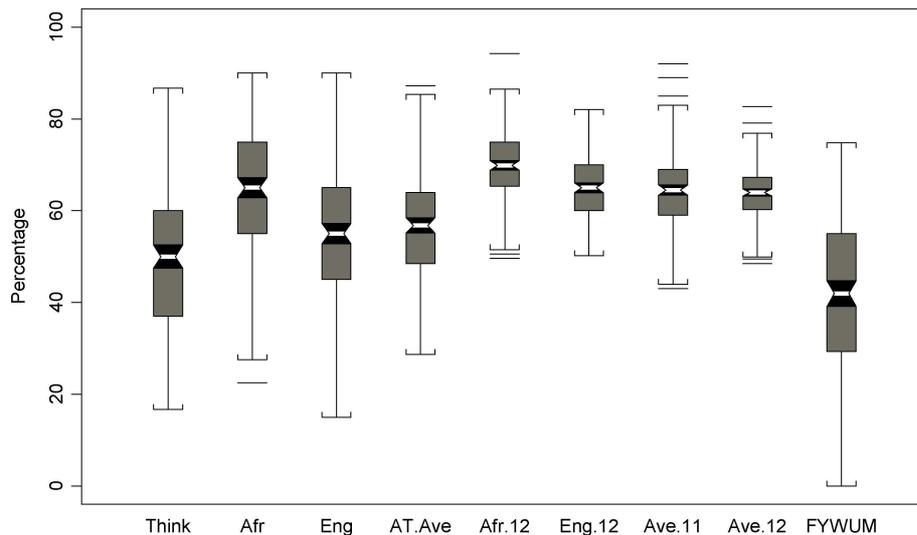
Figure 3: *Notched boxplots of the variables of the Test Battery 3 data set for the 2000 intake groups.*

Figure 3 demonstrates that the confidence interval of the median of FYWUM does not overlap with any of the corresponding confidence intervals of the other variables. Thus, the Test Battery 3 access tests do not give such a good indication of the location of FYWUM as in the case of Test Batteries 1 and 2. A difference of at least 24 percentage points in the means of Ave.12 and FYWUM is observed for each of the three intake groups. These substantial differences are stressed once again in Figure 3, where the upper (third) quartile

of FYWUM is less than the lower (first) quartile of each of the school result variables.

3.2 Correlation matrices associated with school results, access tests and FYWUM variables

3.2.1 Test battery 1

The inter-correlations among the access test, school result and FYWUM variables of Test Battery 1 are displayed in Tables 9 through 11. Note that the critical values appearing at the bottom of each table are the critical values for the hypothesis test, namely that the corresponding population correlation coefficient is significantly larger than zero.

	Maths	Science	Lang	AT. Ave	Maths. 12	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
Maths	1.00	0.49	0.26	0.84	0.11	0.07	0.05	0.08	0.33	**0.09**
Science	0.49	1.00	0.05	0.70	0.02	−0.17	−0.08	−0.05	0.32	**0.10**
Lang	0.26	0.05	1.00	0.60	0.12	0.31	0.49	0.12	0.31	**0.03**
AT.Ave	0.84	0.70	0.60	1.00	0.12	0.10	0.21	0.08	0.44	**0.10**
Maths.12	0.11	0.02	0.12	0.12	1.00	0.08	0.03	0.03	0.29	**0.20**
Afr.12	0.07	−0.17	0.31	0.10	0.08	1.00	0.48	0.48	0.51	**0.15**
Eng.12	0.05	−0.08	0.49	0.21	0.03	0.48	1.00	0.38	0.36	**0.14**
Ave.11	0.08	−0.05	0.12	0.08	0.03	0.48	0.38	1.00	0.45	**0.28**
Ave.12	0.33	0.32	0.31	0.44	0.29	0.51	0.36	0.45	1.00	**0.45**
FYWUM	0.09	0.10	0.03	0.10	0.20	0.15	0.14	0.28	0.45	**1.00**

Table 9: *Inter-correlations among the Test Battery 1 variables – 1999 [Critical r = 0.18 (5% significance level), 0.26 (1% significance level)].*

	Maths	Science	Lang	AT. Ave	Maths. 12	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
Maths	1.00	0.57	0.24	0.81	0.09	0.04	0.11	0.18	0.40	**0.03**
Science	0.57	1.00	0.35	0.85	0.13	0.13	0.20	0.24	0.54	**0.20**
Lang	0.24	0.35	1.00	0.64	−0.08	0.37	0.42	0.14	0.33	**0.04**
AT.Ave	0.81	0.85	0.64	1.00	0.07	0.22	0.30	0.25	0.55	**0.12**
Maths.12	0.09	0.13	−0.08	0.07	1.00	0.07	0.00	−0.11	0.31	**0.22**
Afr.12	0.04	0.13	0.37	0.22	0.07	1.00	0.51	0.37	0.52	**0.23**
Eng.12	0.11	0.20	0.42	0.30	0.00	0.51	1.00	0.31	0.51	**0.14**
Ave.11	0.18	0.24	0.14	0.25	−0.11	0.37	0.31	1.00	0.53	**0.19**
Ave.12	0.40	0.54	0.33	0.55	0.31	0.52	0.51	0.53	1.00	**0.34**
FYWUM	0.03	0.20	0.04	0.12	0.22	0.23	0.14	0.19	0.34	**1.00**

Table 10: *Inter-correlations among the Test Battery 1 variables – 2000 [Critical r = 0.12 (5% significance level), 0.17 (1% significance level)].*

Since the correlations between the university performance and each of the access test and the school result variables are of interest for the purposes of this study, these correlations are bold faced. These tables show that the variable which has the highest correlation with FYWUM is the Grade 12 mark (Ave.12), explaining 20.25%, 11.56% and 44.89% of the variation in FYWUM, for the 1999, 2000 and 2001 intake groups respectively. It is clear that the ability of the access tests to predict university performance is poor in 1999 and

	Maths	Science	Lang	AT. Ave	Maths. 12	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
Maths	1.00	0.62	0.41	0.87	0.44	0.33	0.33	0.51	0.64	**0.40**
Science	0.62	1.00	0.45	0.86	0.45	0.35	0.43	0.55	0.68	**0.44**
Lang	0.41	0.45	1.00	0.71	0.23	0.51	0.50	0.51	0.57	**0.37**
AT.Ave	0.87	0.86	0.71	1.00	0.47	0.47	0.50	0.63	0.77	**0.50**
Maths.12	0.44	0.45	0.23	0.47	1.00	0.26	0.29	0.44	0.60	**0.42**
Afr.12	0.33	0.35	0.51	0.47	0.26	1.00	0.48	0.59	0.62	**0.41**
Eng.12	0.33	0.43	0.50	0.50	0.29	0.48	1.00	0.52	0.58	**0.42**
Ave.11	0.51	0.55	0.51	0.63	0.44	0.59	0.52	1.00	0.84	**0.61**
Ave.12	0.64	0.68	0.57	0.77	0.60	0.62	0.58	0.84	1.00	**0.67**
FYWUM	0.40	0.44	0.37	0.50	0.42	0.41	0.42	0.61	0.67	**1.00**

Table 11: *Inter-correlations among the Test Battery 1 variables – 2001 [Critical $r = 0.07$ (5% significance level), 0.10 (1% significance level)].*

2000, as reflected by the r^2 value, which is never in excess of 0.04. However, the ability of the access tests and school results to explain the variation in FYWUM showed an increase when the threshold for exemption from writing the Test Battery 1 access tests was raised for the 2001 intake group (*effect of the Health Sciences 2001 students*). This suggests that the access Test Battery 1 in its present form does not succeed in predicting FYWUM satisfactory in the case of students with relatively weak school results.

Furthermore, inspection of the correlation matrices reveals that the school result variable Ave.12 is the single variable explaining most of the variation in access test variables Maths, Science and AT.Ave. Moreover, the variable explaining the highest degree of variation in Lang is Eng.12, except for the 2001 intake group, where Ave.12 explains 32.49% of the variation in Lang.

3.2.2 Test battery 2

The correlation matrices associated with the 1999 and 2000 intake groups in the Economic and Management Sciences Faculty are similar to those in Tables 9 and 10 and are thus not reproduced here. The correlation matrix for the 2001 intake group is given in Table 12.

	Maths	Science	Lang	AT. Ave	Maths. 12	Afr. 12	Eng. 12	Ave. 11	Ave. 12	FYWUM
Maths	1.00	0.61	0.29	0.83	0.18	0.16	0.17	0.15	0.42	**0.23**
Numer	0.61	1.00	0.41	0.87	0.24	0.11	0.28	0.28	0.51	**0.34**
Lang	0.29	0.41	1.00	0.67	0.15	0.36	0.39	0.27	0.40	**0.29**
AT.Ave	0.83	0.87	0.67	1.00	0.24	0.24	0.34	0.29	0.56	**0.36**
Maths.12	0.18	0.24	0.15	0.24	1.00	0.12	0.17	0.19	0.47	**0.22**
Afr.12	0.16	0.11	0.36	0.24	0.12	1.00	0.31	0.36	0.46	**0.25**
Eng.12	0.17	0.28	0.39	0.34	0.17	0.31	1.00	0.35	0.53	**0.28**
Ave.11	0.15	0.28	0.27	0.29	0.19	0.36	0.35	1.00	0.65	**0.39**
Ave.12	0.42	0.51	0.40	0.56	0.47	0.46	0.53	0.65	1.00	**0.51**
FYWUM	0.23	0.34	0.29	0.36	0.22	0.25	0.28	0.39	0.51	**1.00**

Table 12: *Inter-correlations among the Test Battery 2 variables – 2001 [Critical $r = 0.11$ (5% significance level), 0.15 (1% significance level)].*

Table 12 reveals that AT.Ave and Numer explain 12.96% and 11.56% of the variation in FYWUM respectively, while Ave.12 is able to explain 26.01% of the variation in FYWUM. Although the correlations of the school result variables with FYWUM are generally greater than the correlations of the access test variables with FYWUM, all these correlations are relatively low. This phenomenon may be attributed to the fact that the access tests of Test Battery 2 were written by prospective students who had Grade 11 and Grade 12 marks mainly between 50% and 70%. This restricted spectrum of students led to a reduction in the lower and upper tails of the distributions of the access test, school result and FYWUM variables. This is in contrast to the *effect of the 2001 Health Sciences students,* which caused a definite increase in the correlations between FYWUM and each of the access test and school result variables (*cf.* Table 11).

The Test Battery 2 correlation matrices depict rather weak linear relationships between the access test and school result variables. It is noted that the variation in Maths and Numer are explained poorly by the school result variables for the 1999, 2000 and 2001 intake groups, but Maths explains the greatest amount of variation in Numer and vice versa: 44.89% in 1999, 33.64% in 2000 and 37.21% in 2001. Eng.12 explains the most of the variation in Lang for the 1999 and 2000 intake groups. In 2001 Ave.12 shows the highest correlation with each of the access test variables among all the school result variables.

3.2.3   Test battery 3

The inter-correlations between the access test, school result and FYWUM variables for Test Battery 3 show similar patterns to those obtained for Test Battery 2 in Table 12. Therefore these correlation matrices are not given here, except to note that Ave.12 shows the highest correlation with FYWUM and explains 18.49% of the variation in FYWUM. The other school result and access test variables, however, have a correlation of less than 0.2 with FYWUM. These low correlations may be explained by the fact that Test Battery 3 access tests were mainly written by prospective students with Grade 11 or Grade 12 averages between 50% and 70%.

Furthermore, as far as the relationships between the access test and school result variables are concerned, it is shown that most of the variation in the access test variable Afr is explained by the school variable Afr.12. Variable Eng.12 explains most of the variation in Eng and Think for the 1999 intake group, while in 2000 and 2001 Eng and Think explain most of the variation in each other respectively.

4   Conclusions

The introduction to this article explained why it is imperative for universities in South Africa to reconsider admission criteria to their institutions. Stellenbosch University developed a battery of access tests to supplement school results in decisions regarding admission. In this article the relationships among results in these access tests, school results and first year university performance are investigated.

Initial explorative analyses were conducted in order to study the univariate properties of

the access test, school result and FYWUM variables, and to compare the properties of these variables. Notched boxplots for three access Test Batteries demonstrated not only that the variations of FYWUM and the access test variables exceeded those of the school result variables, but also the difference in the location of these variables. Indeed, the mean of FYWUM is at least 20 percentage points lower than the corresponding mean of Ave.12 for each of the data sets. Furthermore, the general tendency of the upper (third) quartile of FYWUM to lie below the lower (first) quartile of each of the school result variables (except Maths.12) emphasises that school results create unrealistic expectations of university performance. Whether school results are still a reliable indicator of a prospective student's preparedness for higher education is a matter of great concern. The locations of selected access test variables and FYWUM, however, match closely, suggesting that access test results give a prospective student a more accurate indication of his/her expected average first-year performance at university.

The correlation matrices reveal that particular school result variables have a higher correlation with FYWUM than the access test variables do. Overall Ave.12 is the single variable explaining the largest proportion of the variation in FYWUM. Nonetheless, this proportion of variation explained by Ave.12 is still relatively low. It should be stressed that, regardless of higher correlations between school results and FYWUM, only a small percentage of students included in this investigation obtained a FYWUM of 50% or more. Although the formula used in this investigation for expressing first-year university performance in a single score is not beyond dispute, the differential between FYWUM and school results is a cause for grave concern. These findings indicate that prospective students could find it extremely difficult maintaining their school performance at university.

It was demonstrated that raising the threshold for exemption of writing access tests is accompanied by an increase in the means and medians of access test variables as well as FYWUM (*effect of the Health Sciences 2001 students*). In fact, the mean of FYWUM then exceeds 50% for the first time. Furthermore, this resulted in a substantial increase in the correlations between FYWUM and each of the school result and access test variables, as well as in the inter-correlations among the access test and school result variables.

Although the access tests convey important information regarding the preparedness of a prospective student for university, these tests do not discriminate satisfactorily between *potentially* successful and unsuccessful students. Authorities developing access tests in particular and criteria for admission to higher education in general should be cognisant of the results obtained in this investigation. In the second article [3] our main focus will be on the statistical distributions of the variables reported here. Not only should such analyses confirm the results of the present article, but information crucial for further refinement of admission criteria will be provided.

References

[1] EDUCATION WHITE PAPER 3, 1997, *Education white paper 3: A programme for the transformation of higher education*, Department of Education, Pretoria.

[2] JOHNSON RA & WICHERN DW, 2002, *Applied multivariate statistical analysis*, 5th Edition, Prentice-Hall, Upper Saddle River (NJ).

[3] LE ROUX NJ, BOTHMA A & BOTHA HL, 2004, *Statistical properties of indicators of first-year performance at university*, ORiON, to appear.

[4] NATIONAL PLAN FOR HIGHER EDUCATION, 2001, Ministry of Education, Pretoria.

[5] MCGILL R, TUKEY JW & LARSEN WA, 1978, *Variations of box plots*, The American Statistician, **32**, pp. 12–16.

[6] MITCHELL G, FRIDJHON P & HAUPT J, 1997, *On the relationship between the matriculation examination and university performance in South Africa — 1980–1991*, South African Journal of Science, **93**, pp. 382–387.

[7] VAN WYK JA & CRAWFORD JL, 1984, *Correlation between matric symbols and marks obtained in a first year ancillary physics course at the University of the Witwatersrand*, South African Journal of Science, **80**, pp. 8–9.

[8] VOGEL CA, 1995, *Predictor variables of academic success for first-year technikon students in applied science*, Masters thesis, University of Port Elizabeth, Port Elizabeth.