Foundation Science Student Performance explained. Insight gained at UKZN.

N.F. Kirby[†]

† Centre for Science Access, University of KwaZulu Natal, 3209 Pietermaritzburg, South Africa, email: felthamn@ukzn.ac.za

Alternative Access routes into tertiary science degrees have become a well recognized option for Black South African students with academic potential but who do not make Science Faculty entry requirements (Altink, 1987; Downs, 2005; Downs, 2010; Mabila et al., 2006; Parkinson, 2000; Rollnick, 2006; van der Flier, Thijs & Zaaiman, 2003; Wood & Lithauer, 2005; Zaaiman, 1998). Rollnick (2006) cites Pinto (2001) when claiming that, by 2001, almost every university (including those institutions formerly considered technikons) in South Africa was offering some form of alternative access to disadvantaged students, most of these programmes being science based.

Foundation Programmes in particular have been at the fore-front of educational change in South Africa. Defined by Kloot, Case and Marshall (2008) as “special programmes for students whose prior learning has been adversely affected by educational or social inequalities” (pg.800), these authors report that, the Foundation Programme of the Centre for Science Access (CSA) at the University of KwaZulu-Natal is “widely regarded as an outstanding effort in the genre of foundation programmes (pg. 806)”.

The Foundation Programme of the CSA at UKZN- A Brief History.

The Science Foundation Programme (SFP) was launched in 1991 on the Pietermaritzburg campus of the University of Natal (UN), six years before the Department of Educations’ White Paper (Grayson, 1993; Grayson, 1996). The purpose of this programme was to provide foundational and other relevant courses for previously disadvantaged students who did not meet the formal entrance requirements of the Faculties of Science and Agriculture (Pietermaritzburg), Science (Durban), Engineering and Medicine, but who were judged to have the potential to succeed in degree programmes in those faculties. On the Durban campus of UN a four-year BSc curriculum in which students were admitted directly into first-year courses was also in operation by 1991 (Parkinson, 2000); this has always been referred to as the ‘Augmented Programme.’

Similarly, from 1999 the University of Durban Westville (UDW) operated its own Science Foundation Programme which was quite different from that of UNP. The merger between UN and UDW into the University of KwaZulu-Natal (UKZN) in 2004 saw the amalgamation of the alternative access programmes and the formation of the Centre for Science Access (CSA). The CSA was built upon the premise that students to whom alterative access may be afforded by the Faculty, while all from disadvantaged educational backgrounds, are not a homogeneous group in terms of academic preparedness (Centre for Science Access, 2005). Consequently, the CSA now offers both the Foundation and Augmented Programmes. The former has been operating on both the Pietermaritzburg and Westville campuses since the merger, following the model of the original SFP of UN in terms of educational philosophy, resources and curriculum (Centre for Science Access, 2005). The Augmented Programme moved in 2006 to the Westville campus from Howard College (Durban), and in 2007 was instituted on the Pietermaritzburg campus of UKZN as well.

Notably, the Foundation and Augmented Programmes are **not** bridging courses that were designed to “fill in the gaps left by inadequate schooling” as described by Kloot, Case and Marshall (2008, pg. 801). As these authors point out, bridging programmes have been considered to provide “academic support” as opposed to the “academic development” of ‘foundation programmes’. A central tenet of ‘foundation programmes’ (Ibid) in general is that the disadvantaged students in such a programme need more *time* and more *tuition* in laying the foundation for their mainstream studies (Kloot, Case & Marshall, 2008). These authors claim that the (original) SFP was much more than this, describing it as an “holistic model” (pg. 805); one may be inclined to assume that the current Foundation Programme of the CSA may lay claim to this to, considering that it has changed in name only.

Access to the Foundation Programme of the CSA.

Students are only accepted into the CSA programmes if they have come from disadvantaged schools. These are schools that fall into quintiles 1 to 4 according to the Department of Education “poverty index” based on the physical condition of schools and the poverty of the surrounding community (DOE, 2006). The index is used for resource targeting purposes (where NQ1 is the poorest and NQ5 is the least poor school); NQ1 and NQ2 are nofee schools. Students from quintile 5 schools are usually not considered. These schools in quintile 5 are in most cases ex-Model C schools.

Aside from the "disadvantaged" criterion, various additional stipulations and minimum criteria have been applied over the years as the Centre has grown and contextual factors have changed. There has never been a racial criterion for selection into the Foundation Programme, although the majority of students have tended to be black African with far fewer Indian and Coloured students registering.

Students need at least a National Senior Certificate (or pre-2008 equivalent) including Mathematics, Physical Science and/or Biology or Agricultural Science to be considered for the Access programmes at UKZN. Tables 1 and 2 provide minimum criteria for selection into the Faculty of Science and Agriculture at UKZN prior, and post, to the implementation of the National Senior Certificate (NSC) in 2008. As well as these minimum criteria, the procedure for selection into the CSA Foundation Programme includes the administration of in-house maths and science selection tests. A "selection model score" is calculated for each prospective student based on performance in these tests and school maths and science results (a composite value referred to as "M score"). The selection tests have been undergoing continuous and rigorous investigation over the past 10 years; the current Maths and Science selection tests look little like the original ones and indeed, with every passing year, they continue to evolve (Grussendorf, Liebenberg & Houston, 2004). Only a proportion of those students who apply to the Programme are invited to sit these selection tests; a small minority of those who do, are offered places in the Programme. For example in 2009, 1161 students applied for the Access Foundation streams, 63% of whom met the minimum criteria and wrote the tests. Only 23% of those who wrote the tests met the selection test criteria. In other words, 14% of students who applied to the Programme were offered places (Yvette Chetty, CSA Selection Officer, personal communication, March, 2009). Furthermore, a number of those students offered places, do not enter the Programme for various reasons (e.g. financial or personal). Students entering the Programme with a matriculation endorsement enter the BSc 4 Foundation stream of the Foundation Programme; those without endorsement are registered in the SFP stream. This distinction is made only for the purpose of carrying credits into their degrees (the former do, the SFP do not). In practice though, Foundation students experience exactly the curriculum in their Access year, and indeed are in the same class groups.

An in-house language test has also been used in the past to select foundation students, but this has in recent years been replaced by the Standardised Assessment Test for Access and Placement (SATAP) - English for Academic Purposes Test, and has been used for placement into the different academic literacy modules that Access students are obliged to take, rather than for selection (Parkinson, 2005).

Table 1
Criteria for selection into the Faculty of Science and Agriculture, UKZN prior to 2009.

	Senior Certificate (Matric) level (minimum)			
Student group	Endorsement required	APS	Maths	Science *
Foundation Programme				
SFP	None	20	SG F	SG F
BSc 4 Foundation stream	Full	20	SG F	SG F
Augmented Programme (BSc 4 Augmented stream)	Full	28	HG E *or* SG B	HG E *or* SG B
Direct Entry **	Full	32	HG E *or* SG B	HG E *or* SG B

Notes. Criteria given as laid out in the Faculty of Science and Agriculture, UKZN handbooks for respective years. SG = standard grade, HG = higher grade.
*Science subject may be either Biology, Physical Science or Agricultural Science.
** For entry into the general BSc LES undergraduate degree
Appendices A and B describe the values of the symbols and levels of performance respectively, and the calculations for the Admission Points Score (APS), a composite score, also used for admission.

Table 2
Criteria for selection into the Faculty of Science and Agriculture, UKZN from 2009 onwards

	National Senior Certificate (NSC) (minimum)				
Student group	Admission to Bachelor's degree	APS	English *	Maths	Science **
Foundation Programme					
SFP	No- NSC	16	Level 4	Level 2	Level 2
BSc 4 Foundation stream	Yes NSC (Deg)	16	Level 4	Level 2	Level 2
BSc 4 Augmented stream	Yes NSC (Deg)	22	Level 4	Level 3	Level 3
Direct Entry ***	Yes NSC (Deg)	28	Level 4	Level 4	Level 4

Notes. Criteria given as laid out in the Faculty of Science and Agriculture, UKZN handbooks for 2009 and Admissions Policy (2008). For those students who matriculated before 2008, criteria in Table 1 apply.
* Life Orientation level 4 also required.
**Science subject may be either Life Science (Biology), Physical Science or Agricultural Science.
***For entry into the general BSc LES undergraduate degree.
The transition from a system of grades and symbols to one of levels has required some normalization. Appendix C provides an indication of parity across the two systems (Umalusi, 2010).

The Foundation Programme.

Foundation students enrol in a stand-alone composite curriculum consisting of compulsory, year-long foundation maths, physics, chemistry and biology modules. They also have to complete two academic literacy modules. These, and a counselling component are integrated into the timetable of the foundation "package". There are no elective modules. The subjects run concurrently throughout the year to maximize the opportunities for transfer of knowledge between the subjects.

The Foundation Programme curriculum differs from a bridging programme in that it does not assume that the students enter at a level close to what is needed for entrance into the University environment, but assumes that students need to build a foundation for meaningful learning, in many cases for the first time. Students are not pre-taught for their undergraduate degree as what is important is that they acquire flexible, transferable learning strategies and appropriate study habits rather than familiarize themselves with content from first year courses. This foundation is built in a phased transmission manner, where the beginning and end of the programme are matched to where the students come from (first semester) and where they wish to go next (second semester, and subsequently into the University mainstream). This transition is phased in terms of pace of work, quantity of work, scaffolding required and level of difficulty, the intention being that students are able to operate in their (Vygotskian) zone of proximal development.

The broad integration of disciplines is an important aspect of the Foundation Programme curriculum; a demonstration of the unity (and diversity) of science to students is thus made possible. Transfer of learning from one context to another has been made possible by the mutual development of the curriculum by an inter-disciplinary team and also by the continuous discussion amongst members of the staff from each discipline about transfer opportunities. Indeed, Trowler (2008) from interviewing the CSA staff on the Pietermaritzburg campus has acknowledged the Foundation Programme staff as a fine example of a community of practice. Certainly this has value for effectively delivering a curriculum that the Foundation Programme aspires to do.

Accountability of the CSA.

Periodic reviews (mediated by the University's Quality Promotion Unit) have required the CSA to account for its philosophy, curriculum, actions and decisions, and student performance, to the institutional community (Southway-Ajulu, 2005), and also to broader society, including the private donors from whose generous funding the Centre has benefited. Furthermore, the South African Government's funding policies (DOE, 2006b) have resulted in student performance and throughput being even more closely scrutinized by Faculty (who is accountable to the Government) than has been the case in the pre-merger access programme initiatives. On the basis of student enrolment and through-put, government funding is either extended or withdrawn. Student retention is of particular concern in the Foundation Programme since, to avoid being excluded from the Science Faculty of UKZN, students must pass every one of their five requisite modules in their access year to continue into mainstream. Those who are lost to the Faculty of Science and Agriculture are very often recommended to other Faculties within the University, to a University of Technology or to a Further Education and Training (FET) facility. However, because of their weak entrance qualifications that placed them in the Foundation stream in the first place, they are unlikely to be admitted to any other university science faculty. Indeed, records show that only a proportion of Foundation Programme students pass their access year (Table 3).

Table 3
Number of students proceeding from the Foundation Programme (2006-2009)

Year	Original intake	Attrition from Programme *[1]	Number of students proceed from Programme
2009	109	64	45
2008	87	42	45
2007	66	41	25
2006	97	62	35

Note. *1 Attrition includes a few (around 1 to 3) students who withdraw each year for a variety of personal reasons, the balance are academically excluded either in June or November.
Results reflect proceed rates after the release of supplementary exams results.

Previous research has shown that students who have accessed mainstream studies by successfully completing the Foundation Programme have gone on to perform well (Downs, 2005, 2010; Kirby, in prep). Downs (2010) concludes that, whilst also appearing to be an effective system for selection into mainstream, the Programme has significantly increased the quantity and quality of science graduates in South Africa and highlights the valuable role played by such a programme in issues of access and redress.

However, these high rates of exclusion at the end of the Foundation year are severely problematic for both the CSA and the Faculty. This study is a response to this need to increase the numbers of students successfully completing the Foundation year so that they can go on to fulfil their potential at mainstream level. In gaining insight into the factors affecting performance of the Foundation Programme students in their access year, the challenges and opportunities for improvement may be better understood. It is a chance to examine the selection processes at a time when the large majority of students who apply to the Foundation Programme are turned away because they do not meet the criteria for selection. Most importantly, identifying possible opportunities for remediation in the Programme and its existing curriculum will contribute to maximizing a student's potential, and their preparedness for successive study. This certainly merits investigation given the national imperative to extend access to tertiary education, and in particular, to the Sciences. Classification and regression tree (CRT) analysis has been found to be most useful in this regard.

Methodology: Classification and regression tree analysis.

"Research approaches that try to isolate the influences of a few variables for all students will simply miss the point and probably provide little in the way of useful, practical or policy relevant evidence" (Pascarella & Terenzini, 1998, pg. 155). Heeding this concern, classification and regression trees (CRT) of Breiman, Friedman, Olshen, and Stone (1984) are an attractive non-parametric alternative to generalized linear modelling techniques conventionally used. Trees may also be used as a descriptive and exploratory technique to support traditional regression models. The main difference between classification and regression trees is that the response variable described by the former is categorical, the latter refers to scale, continuous data (Breiman et al., 1984; De'ath & Fabricius, 2000; El-Emam, Goldenson, McCurley & Herbsleb, 2001). Explanatory variables can be nominal, ordinal or continuous for both classification and regression trees.

The computer package SPSS Base (version 15 for Windows) offers the Classification Trees™ procedure, with one of the available growing methods being CRT (SPSS Inc, ver. 15, 2006). Analysis for this research was conducted using this CRT option. The Classification Tree procedure can be used for "segmentation and stratification" (identifying cases that are likely to be members of a particular group, or assigning them into one of several categories respectively), "prediction" (to create a rule for the prediction of future events such as a student passing or failing, or achieving a particular mark), "data reduction and variable screening" (to select a useful subset of variables from a larger set to describe and explain an outcome/response variable) and "interaction identification" (identifying relationships that refer to only specific subgroups of, for example, students) (SPSS Inc, 2004, pg.2).

Not only does tree analysis avoid the complexities and restrictive assumptions of logistic and non binomial regression modelling, but they have particularly clear, visual appeal, and are easy to understand and interpret. Classification and regression trees are built by using a binary partitioning algorithm to recursively divide data into relatively homogenous, dichotomous groups, thus revealing the explanatory variables which best describe the response variable (Breiman et al., 1984). The analysis exposes a hierarchy of context dependent effects of the explanatory variables, which allows a clear picture of the interaction between factors influencing the response variable, to emerge.

With data being recursively partitioned (a "parent" into two "child" nodes) an attempt is made to maximize within-node homogeneity. Recursive splitting results in cases (students) being classified into smaller and smaller nodes, the similarity in the outcome variable within each node increases at the same time. Similarly, the difference in the outcome variable between nodes also increases. The extent to which a node represents a heterogeneous subset of cases (in this research, students) is an indication of "impurity". This is measured by the least-squared deviation (LSD) measure of impurity for continuous outcome variables. A terminal node in which all cases have the same outcome value is regarded as being "pure". For nominal and ordinal outcomes (categorical data), a number of measures of impurity exist (SPSS Inc, 2004). The Gini impurity measure was selected for this analysis. Here, splits are based on squared probabilities of membership for each category of the outcome variable with the view to maximizing homogeneity in each child node. A reduction in impurity can be calculated by comparing impurity of the root node with the sum of the impurities of the child nodes (Breiman et al., 1984).

For each split, each explanatory variable is evaluated to find the best cut point (continuous data) or groupings of category (nominal or ordinal outcomes). The explanatory variable that yields the large reduction in impurity is chosen for the first split (Breiman et al., 1984). "Improvement", indicated on resulting trees, refers to the improvement in purity of child nodes resulting from a split of the parent node by the explanatory variable used to make the partition. The pruning criterion was applied to avoid over-fitting of the models (SPSS Inc, 2004). This means that after a tree is grown to its full depth (until the stopping criteria are met), pruning trims it down to the smallest sub-tree that has an acceptable risk value.

CRT can use surrogates for explanatory variables where values for particular cases may be missing and where a high association with the original variable exists. Surrogate splitters are explanatory variables that are not as good at splitting a group as the primary splitter but which yield similar splitting results; they mimic the splits produced by the primary splitter. Examination of surrogates and alternative splits can lead to a more complete understanding of competing explanatory variables and their relationships (De'ath & Fabricius, 2000). The association measure (the lambda coefficient for contingency tables, λ) indicates the degree to which splits based on the surrogate match those based on the actual predictor. The largest possible association value is 1.0 which means the surrogate mirrors the action of the primary splitter in the resulting tree and is a perfect substitute for it; these variables can be used interchangeably (El-Emam et al., 2001). For each variable when it appears as a surrogate, the improvements in purity, had that variable been selected for the primary split are summed up for all nodes. These summed improvements are scaled relative to the best performing variable where the highest value is 100. Thus each explanatory variable's "importance scores" incorporate information both on the use of the variable as primary splitters, in addition to their relative worth as surrogates should the primary splitter be missing. These are ranked (in order of importance to the overall construction of the tree) (Breiman et al. 1985; El-Emam et al., 2001; SPSS Inc, 2004).

Applying a model to other data files containing similar variables to generate the predicted outcome values for each case in that file is referred to as "scoring" (SPSS Inc, 2004, pg 99). In the form of SPSS command syntax, a generated model specifies the "rules" for assigning predicted values to cases in a data set.

CRT models are commonly used in a wide range of fields from medical diagnostics to ecological studies (see Morris and Fynn (2003) for a South African example), and even accident analysis and prevention (e.g. Elmitiny, Yan, Radwan, Russo, & Nashar, 2010.). Hayden, Hayden and Gamst (2004) have used regression trees to identify predictors of success as Emergency Medicine residents from a set a variables available at the time applicants were screened. Although trees are increasingly being used by educational researchers internationally (for an early example, see Grayson, 1997), they appear to have had limited exposure in South Africa. In one South African study by Lourens and Smit (2003), classification trees revealed that students studying in certain "subject matter categories" had a much better chance of being a successful first year student (passing all requisite modules) than others, and success could be predicted on the basis of this category and grade 12 aggregate.

Key work informing this current research is that of Ma (2005). This author claims to have "pioneered the application of CART (an alternative acronym for classification and regression tree analysis) in education research" (pg.86). Indeed as this author says, one of the most attractive features of this tool is its ability to "identify local interactions" that "holds great promises for informing education policy and practice" (pg. 86).

For the current research, two Foundation student cohorts were studied. Those enrolled in the programme in 2008 had left secondary school having written the Senior Certificate. In the 2009 Foundation cohort, some of the students had matriculated before 2008, the majority having written the National Senior Certificate implemented at the end of 2008. This provided the opportunity to compare the influence of the two national schooling systems. A comprehensive set of explanatory variables was included in the analysis of student performance (as indicated by a final mark for each Foundation module for the 2008 and 2009 cohorts, average, overall Foundation mark and proceed/exclude status). Only student performance as indicated by the average overall mark and proceed/fail decision is presented here. Performance in the individual Foundation modules is only alluded to in this paper and is presented elsewhere in detail (Kirby, in prep.).

Biographical variables included students' gender and home language. School history data included school quintile, "matric" points (APS) earned, and their school-leaving performance in Maths, Physics, Biology and English (for the 2009 cohort this was referred to as "APS equivalent" to accommodate the different systems of grades and symbols before, and levels, after the implementation of the NSC). Student performance in the individual selection tests, and the composite selection model score were included to cover issues of selection. The influence of the "M score" (in itself, and not as a component of the selection model score) was secondarily explored as a possible surrogate for the more comprehensive selection model score. Student performance in the Standardised Assessment Test for Access and Placement Test (SATAP) (English for Academic Purposes) was also included in analysis. Socioeconomic data pertained to Foundation students' accommodation and travel arrangements, and the extent of their financial support during their access year. Thus the relative influence of a total of 21 and 20 variables on student performance in 2008 and 2009 respectively was explored (socio economic data in 2009 was not directly collected from students in the form of a questionnaire as was done in 2008 and consequently did not include travel arrangements; accommodation data in 2009 was also simplified).

Student Performance in the Foundation Programme as indicated by Overall Average.

Overall student performance in 2008, as indicated by an average mark of all five foundation modules, can neatly be described by the selection model score (Figure 1). Homogeneity in the daughter nodes was increased by 30% by this one variable, for which the only reasonable surrogate to be found was school Physical Science APS (Improvement in purity = 10.25, λ coefficient for contingency tables = 0.43). It appears that performance in school physical science has, in the past, been a fairly reliable indicator of potential to perform well in the Foundation Programme. Indeed, the influence of this school subject in 2008 has been noted in the individual Foundation Biology, Chemistry and Physics modules (described elsewhere). This variable was ranked second to the selection model score in terms of overall importance in the construction of the tree, the third variable being English SATAP test scores, followed, in fourth place, by financial support. When considered, the "M score" was more effective than these latter two variables in reducing heterogeneity in the root node, but was only half as good as the selection model score (Improvement in purity = 8.68, λ coefficient for contingency tables = 0.54). However, when testing the influence of the "M score" in the absence of the selection model score, this variable and English SATAP scores were shown to be the most important primary splitters, and ranked first and second in overall importance (Figure 2). The heterogeneity of node 1 was reduced by 36% by the "M score", with those achieving fewer than 45 Matric points or less in their science and maths school subjects, performing relatively poorly in the Foundation year (node 3, M = 51.83, SD = 5.17). Furthermore the Maths selection test was revealed as an effective indicator of performance in the group of students who achieved more than a total of 45 points for their maths and science school subjects.

Those students who achieved more than 56 points in the selection model score in 2008 did very well in the Foundation Programme in comparison to those who achieved this score or below (M = 61.89, SD = 6.28; M = 53.23, SD = 6.33 respectively), $t(77)$ =5.73, $p < 0.001$; this being a large effect as described by Field (2009), r = 0.57 (two-tailed). This is mirrored in the 2009 tree for overall average (Figure 3), the value of 56 being reflected again as the cut off score (M = 59.15, SD = 8.99; M = 51.14, SD = 7.07 respectively), $t(86)$ =4.67, $p < 0.001$; r = 0.44 (two-tailed). What is particularly important to note is that the selection model score for automatic selection into the 2008 and 2009 Programmes was 52 selection points. Furthermore, 32 % (in 2008) and 23% (in 2009) of those students admitted into the Programme did not actually achieve this score but were admitted anyway on other grounds. In fact, using the full suite of selection criteria, of which the selection model score is but one criterion (see above), only 27% and 15% received automatic acceptance into the Programme in 2008 and 2009, the balance failing some minimum criterion and being accepted later after consideration by the selection officer and the Dean of Faculty. As effective as the selection model score appears to be at identifying students with potential to succeed in the Programme, it appears it is not being used to its' full potential.

The 2009 tree describing overall student performance (as indicated by average mark for all foundation modules) is more complex than that described for the 2008 cohort (Figure 3). The only variable that considerably reduced heterogeneity in the root node, other than the selection model score was accommodation (Improvement in purity = 10.96, λ coefficient for contingency tables = 0.22). The importance of the maths selection test, already alluded to in 2008, came to the fore in this tree with the small group of overall best-performing students, having achieved more than 77% in this test (node 6). This variable was also ranked second to the selection model score in 2009 in terms of overall importance to the construction of the tree. The "M score" was not a good substitute for the selection model score in 2009 (Improvement in purity = 1.65, λ coefficient for contingency tables = 0.33); furthermore when the selection model score was excluded from the list of variables to generate the 2009 tree, the "M score" did not feature at all as it had done so in 2008.

What is interesting in 2009, is the tendency for students with higher levels of English language proficiency to perform poorly overall in the Foundation modules (node 10, M = 50.75, SD = 5.84). Another inverse relationship revealed here is somewhat more alarming: those weaker students (as indicated by the selection model score) who scored well in the science selection test do particularly badly overall in the Programme (node 8, M = 44.66, SD = 5.03). Given the absence to date of any indication that the science selection test is at all useful in explaining or predicting performance in any of the Foundation modules (Kirby, in prep b), the use of this selection test in the future should be seriously considered.

The positive influence of financial support, particularly for weaker students is shown in this tree, with those students who scored lower on the selection model failing on average, if they do not receive financial support (node 4, M = 44.92, SD = 5.56). This variable, and accommodation (although not revealed as a primary splitter) were ranked third and fourth respectively (out of the 20 variables included) in the overall construction of the tree. These socioeconomic variables have been shown to have had particular influence on the all of the Foundation modules examined independently (Kirby, in prep b).

Given the explanation of overall student performance as described by these trees, it is not surprising that using only the selection model score to generate the rules for tree construction, it was possible to accurately predict student performance in 2009 using this model (predicted and actual average mark for 2009, $r = 0.45$, $p < 0.01$), this being considered a medium to large effect (Field, 2009). The isolated "M score", in conjunction with the individual English SATAP and maths selection test scores was shown to have no value in predicting the 2009 average mark (predicted and actual average mark for 2009, $r = 0.03$, $p < 0.05$), in spite of the apparent ability of these variables to describe the 2008 performance.

With the reiteration of the dominance of the selection model score across the "hard science" Foundation modules (Physics, Maths and Chemistry modules explained elsewhere; Kirby, in prep b), and as a predictor of overall performance across all five modules, it is clear that the selection model score has great value in determining whether a prospective student has potential to succeed in the Foundation Programme or not. Clearly though this model has not been used to its best potential to ensure success after access, as the students have, to date been accepted with scores much lower than 56, the score repeatedly shown to have more discriminatory value. Furthermore, whilst it may have been possible in the past to use the "M score" to discriminate students with the potential to succeed in the Foundation Programme as a whole from those who did not, the value of this score on its own diminished in 2009. A notable exception to these trends has been found in the Foundation Biology module (Kirby, in prep b). Not only has the selection model score been shown to have limited value in explaining success in this module, but specific factors shown to have little value in explaining performance in the 2009 maths, chemistry and physics modules, and overall average this year, namely the "M score" and English SATAP test results, have been shown to have significant value in predicting performance in the Foundation Biology module. This suggests that, selection for, and performance in, the Foundation Biology module is at odds with the other Foundation modules, and with performance in the Programme overall. This clearly has implications for teaching, learning and curriculum development in this particular module.

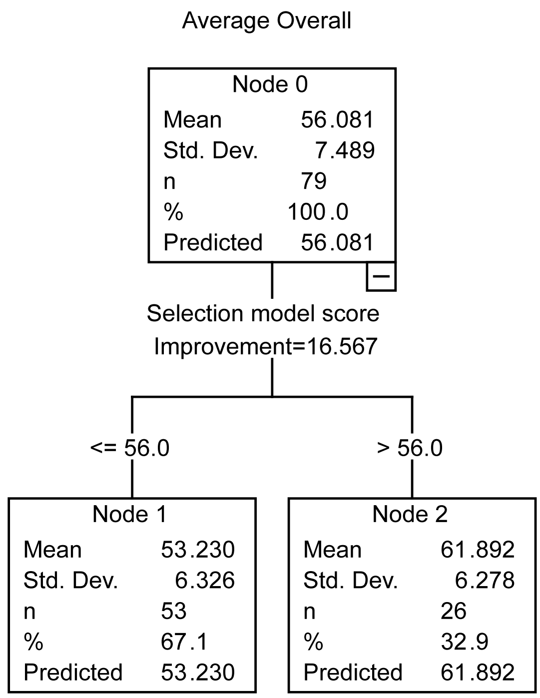


Figure 1. Regression tree for 2008 Overall Final Mark Average (N = 79).

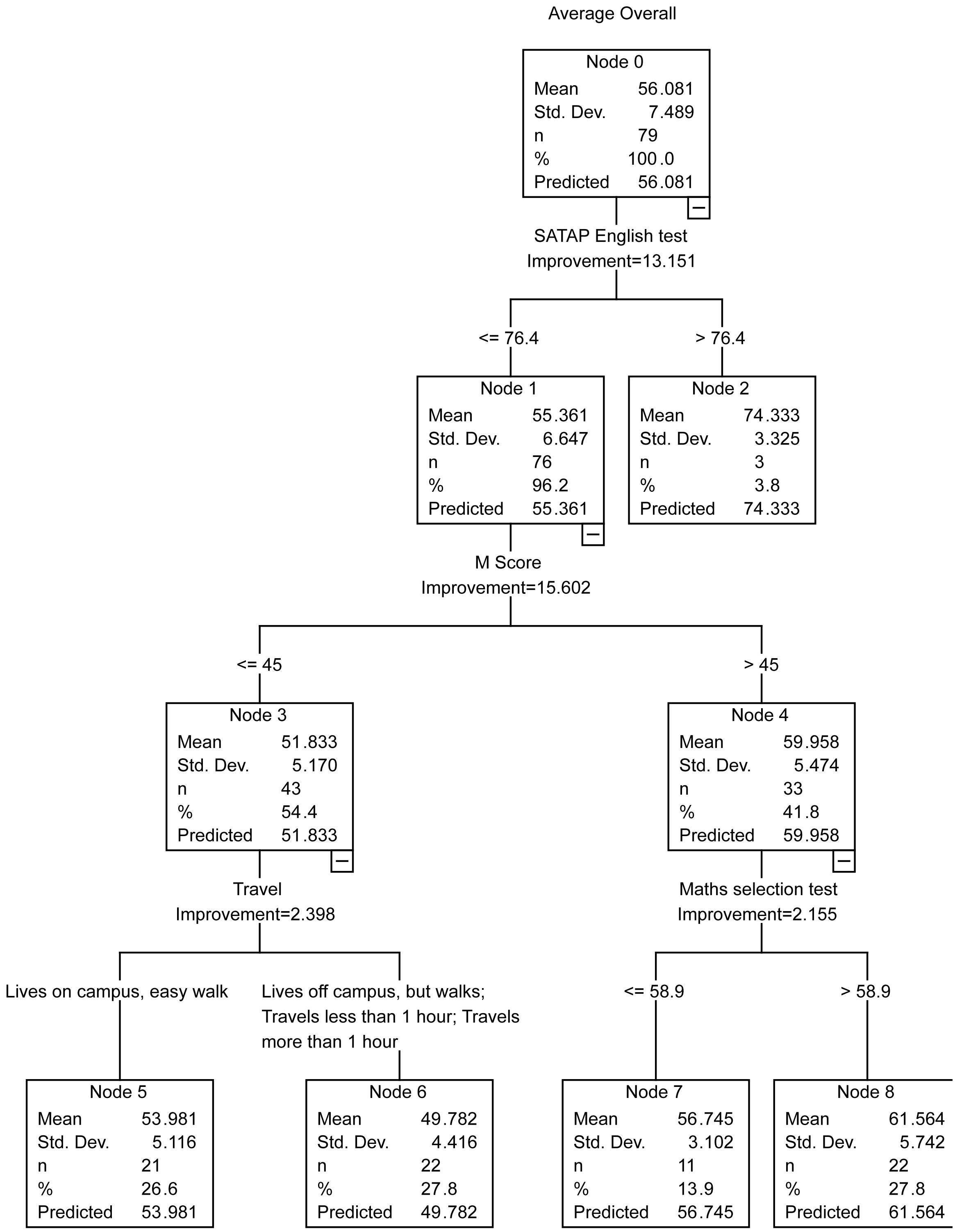


Figure 2. Regression tree for 2008 Overall Final Mark Average (N = 79). The Selection Model score has been replaced by constituent “M score” in the construction of this tree.

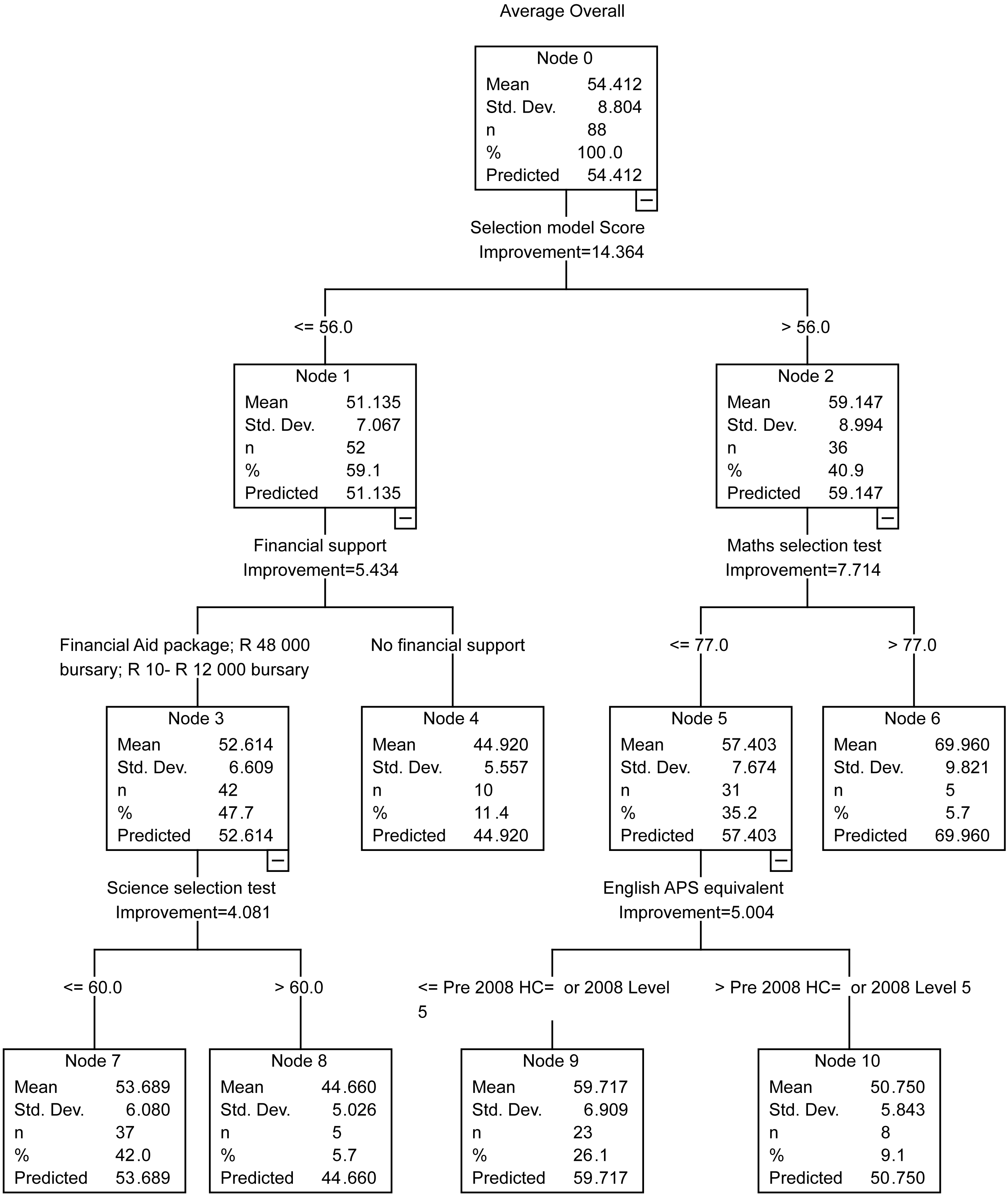


Figure 3. Regression tree for 2009 Overall Final Marks Average (N = 88).

Student Performance in the Foundation Programme as indicated by Proceed Decision.

The classification trees generated for the proceed decision in 2008 and 2009 are similar in number of respects. Most striking is the role played by accommodation; none of the students who rented accommodation in 2008 passed the year (Figure 4), and 25 of the 35 students without a place in residence in 2009 did not proceed (Figure 5, node 1). For those students who did have a place in a University residence (or lived at home in 2008), the selection model score was a very important indicator of their ability to pass the Programme. In both cohorts, the cut off value to best ensure a student proceeded was 50.8 (51) selection model points. A total of only 3 students who achieved less than this score (and had more secure accommodation) passed the Programme over the two year period (Figure 4, node 3). Of those in University residence in 2009, none who achieved less than this score, proceeded from the Programme (Figure 5, node 3).

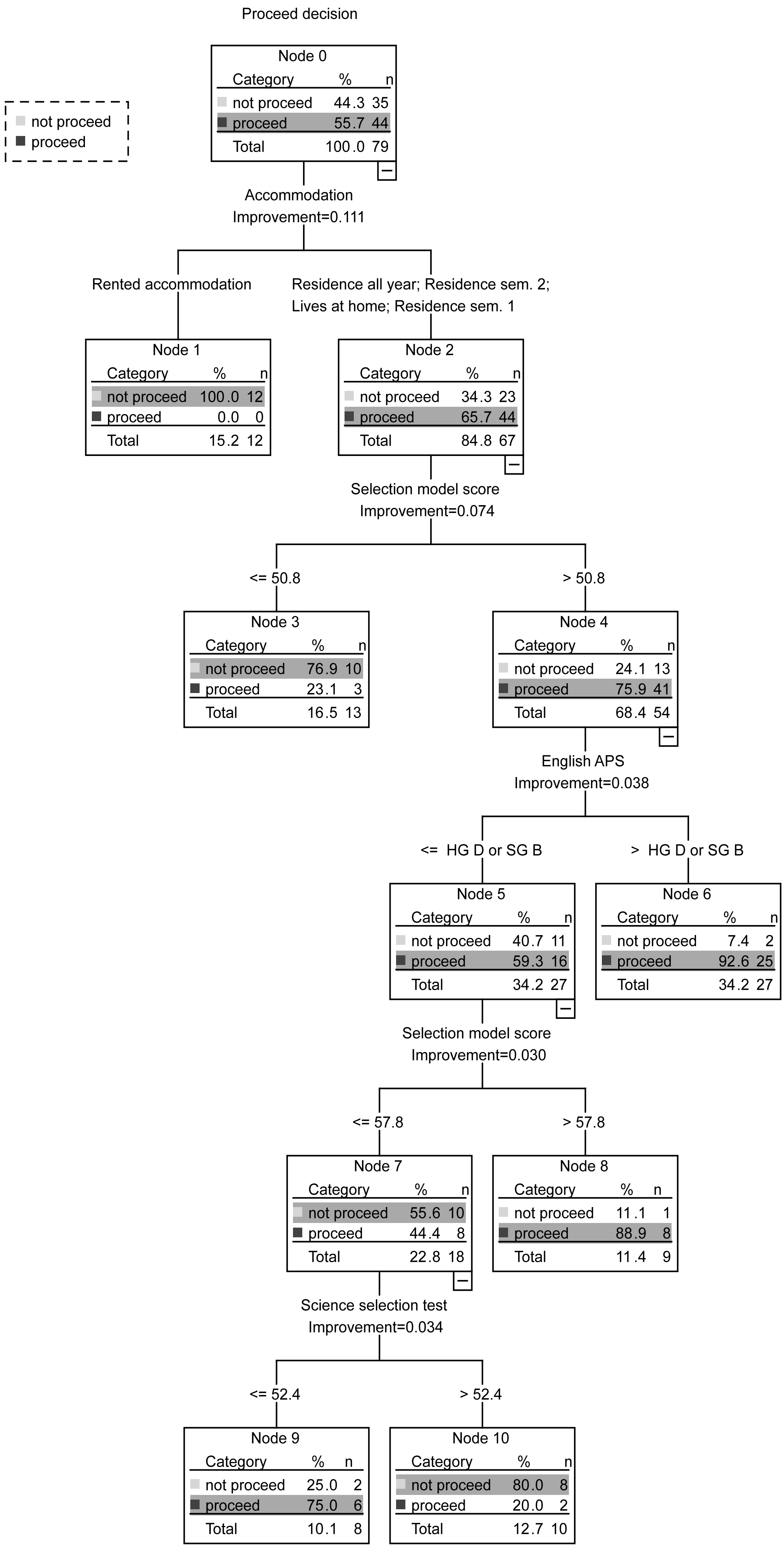
Although accommodation was the primary splitter of both root nodes, it was the selection model score in 2008 and 2009 that was ranked first in terms of overall tree construction, with accommodation in second position in both instances (74.6% and 46.2% relative normalized importance in 2008 and 2009 respectively). This is an indication of the increasing importance of the selection model score as a whole, and the maths selection test component in particular in 2009. In 2009, the constituent maths selection test was revealed as a better surrogate for the selection model score than the isolated "M score" (Improvement in purity = 0.032, λ coefficient for contingency tables = 0.50; Improvement in purity = 0.021, λ coefficient for contingency tables = 0.33 respectively). Furthermore, the selection model score was able to reduce heterogeneity in the 2009 root node almost as well as accommodation (Improvement in purity = 0.059) (this was not so in 2008). What can perhaps be inferred here, is that careful selection of those students leaving school with a National Senior Certificate is increasingly important as the "M score" appears to have lost some of its discriminatory value . Once selected on the basis of some alternative selection mechanism, providing a place in residence will improve that student's chances of proceeding from the Programme considerably.

The effect of implementing the level 4 English requirement in 2009 is also apparent in the trees. In 2008, prior to this being an admission requirement, English APS was the primary splitter of node 4, that is the bulk of students who stood a reasonable chance at proceeding since they had been given a place in residence and had achieved at least 51 points in the selection model. Only 2 students with higher levels of English language proficiency did not proceed (node 6). In 2009, English language proficiency did not feature as a primary splitter and ranked low in terms of relative importance in overall tree construction (15th out of 20 variables). Instead in 2009, the influence of performance in school Physical Science was apparent, with almost all students in node 6 achieving more than a minimum of level 2 in this subject, going on to proceed from the Programme. Although this variable had a relatively small effect on proceed rates in comparison to the selection model score (23.6% relative importance), in context of its role in explaining overall average Foundation marks (as suggested above), it appears that, of all the school subjects included in this analysis, physical science is the most reliable.

The inverse relationship between the science selection test and performance was reiterated in the 2008 proceed/exclude tree (nodes 9 and 10), with those students not scoring well in this test going on to proceed from the Programme and vice versa. This selection test did not feature in the 2009 tree as a primary splitter or as a reasonable surrogate.

Although actual numbers are reflected in these trees, they may also be interpreted in terms of predictive value. The grey highlighted bands in each node serves as an indication of what would be predicted were the trees to be used for predictive purposes. For example, the 2009 model would predict that all students not provided with a place in residence would not proceed. Both the 2008 and 2009 trees had good predictive power: 87.3% and 80.7% respectively. In other words ten students (12.7%) would have been misclassified by the 2008 tree: of the 44 students that proceeded, five would have predicted to fail (88.6% would have been correctly predicted). Similarly, of the 35 who did not proceed in 2008, five would have been predicted to pass (85.7% correctly predicted). The 2009 model was found to have slightly less predictive power. Of the 43 students who did not proceed that year, the model would have incorrectly predicted 7 to pass (83.7% of not-proceeds correct). Ten of those who did actually pass would have been predicted to fail (22.2% incorrectly predicted).

The ability of these models to *explain* proceed rates is very clear; more over their ability to *predict* failure and success is considerable, given the low degree of misclassification as described above. No single school history indicator of potential to successfully proceed from the Foundation Programme appears to exist; furthermore the composite school science and maths scores ("M scores") are insufficient. Combined with alternative selection tests to form the selection model score however, this "M score" can contribute to a powerful selection tool that efficiently discriminates between those students who have the potential to proceed from the Programme, from those who do not. With this confirmed, it would be fool hardy to accept students into the Programme with less than the minimum prescribed selection model point score. More over, it must be acknowledged that there is room to improve this selection model, specifically to reconsider the use of the science selection test. Given the reiteration also of the increasingly important roles accommodation and financial support have been shown to play in the success of 2008 and 2009 students, the mechanisms allowing access to these students, should be extended to these socio-economic issues. This is particularly true of those students who enter already at risk since their academic performance, to date, is weaker. Once granted access to the Programme, it appears that success can be better ensured by investing in the students by providing them with financial support and places in University residence. Having identified problems and areas of potential succeed in



Proceed decision
not proceed
proceed
Node 0
Category % n
not proceed 44.3 35
proceed 55.7 44
Total 100.0 79
Accommodation
Improvement=0.111
Rented accommodation
Residence all year; Residence sem. 2; Lives at home; Residence sem. 1
Node 1
Category % n
not proceed 100.0 12
proceed 0.0 0
Total 15.2 12
Node 2
Category % n
not proceed 34.3 23
proceed 65.7 44
Total 84.8 67
Selection model score
Improvement=0.074
<= 50.8
> 50.8
Node 3
Category % n
not proceed 76.9 10
proceed 23.1 3
Total 16.5 13
Node 4
Category % n
not proceed 24.1 13
proceed 75.9 41
Total 68.4 54
English APS
Improvement=0.038
<= HG D or SG B
> HG D or SG B
Node 5
Category % n
not proceed 40.7 11
proceed 59.3 16
Total 34.2 27
Node 6
Category % n
not proceed 7.4 2
proceed 92.6 25
Total 34.2 27
Selection model score
Improvement=0.030
<= 57.8
> 57.8
Node 7
Category % n
not proceed 55.6 10
proceed 44.4 8
Total 22.8 18
Node 8
Category % n
not proceed 11.1 1
proceed 88.9 8
Total 11.4 9
Science selection test
Improvement=0.034
<= 52.4
> 52.4
Node 9
Category % n
not proceed 25.0 2
proceed 75.0 6
Total 10.1 8
Node 10
Category % n
not proceed 80.0 8
proceed 20.0 2
Total 12.7 10

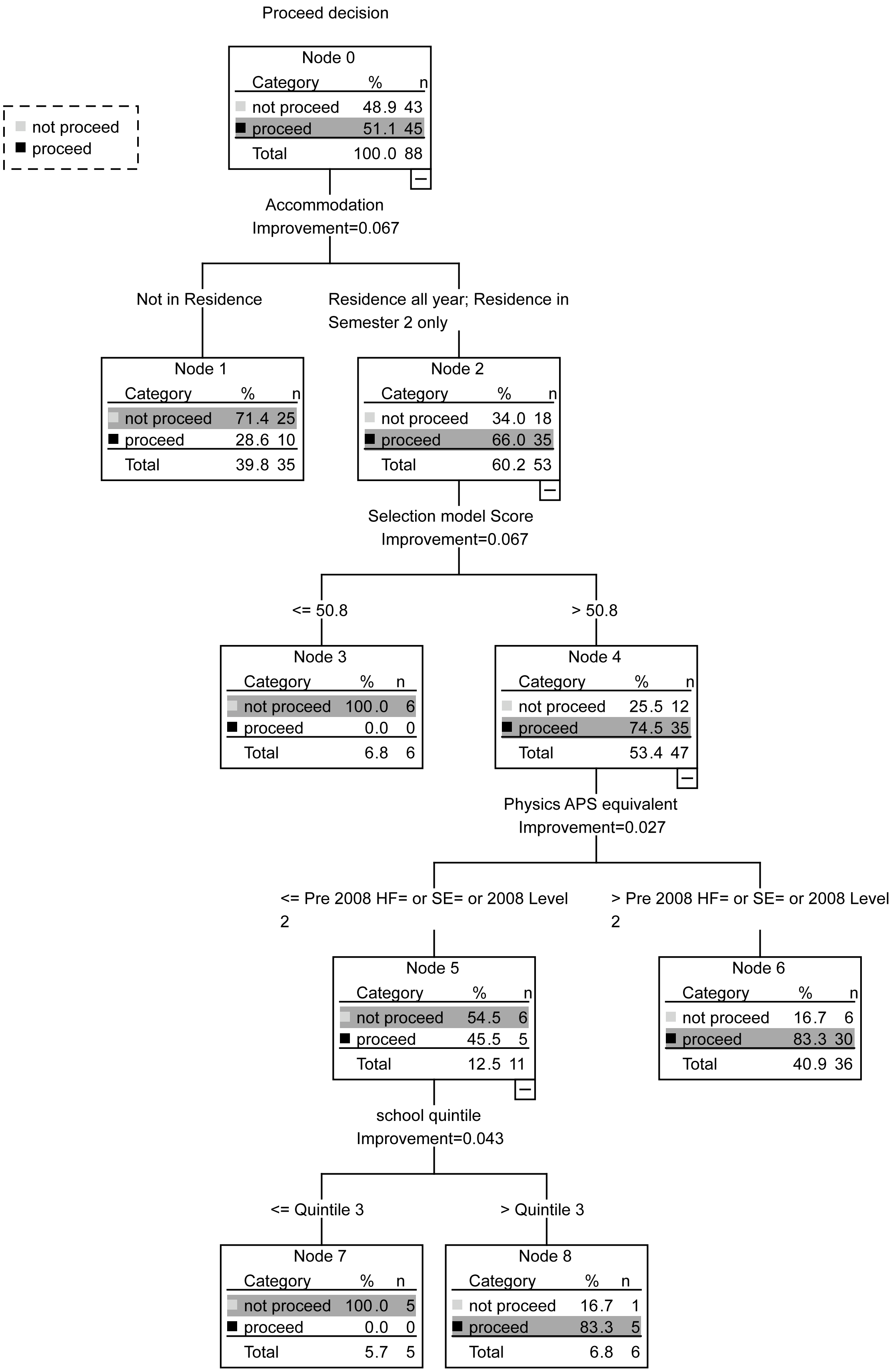

Proceed decision
not proceed
proceed
Node 0
Category % n
not proceed 48.9 43
proceed 51.1 45
Total 100.0 88
Accommodation
Improvement=0.067
Not in Residence
Residence all year; Residence in Semester 2 only
Node 1
Category % n
not proceed 71.4 25
proceed 28.6 10
Total 39.8 35
Node 2
Category % n
not proceed 34.0 18
proceed 66.0 35
Total 60.2 53
Selection model Score
Improvement=0.067
<= 50.8
> 50.8
Node 3
Category % n
not proceed 100.0 6
proceed 0.0 0
Total 6.8 6
Node 4
Category % n
not proceed 25.5 12
proceed 74.5 35
Total 53.4 47
Physics APS equivalent
Improvement=0.027
<= Pre 2008 HF= or SE= or 2008 Level 2
> Pre 2008 HF= or SE= or 2008 Level 2
Node 5
Category % n
not proceed 54.5 6
proceed 45.5 5
Total 12.5 11
Node 6
Category % n
not proceed 16.7 6
proceed 83.3 30
Total 40.9 36
school quintile
Improvement=0.043
<= Quintile 3
> Quintile 3
Node 7
Category % n
not proceed 100.0 5
proceed 0.0 0
Total 5.7 5
Node 8
Category % n
not proceed 16.7 1
proceed 83.3 5
Total 6.8 6

Figure 5. Classification tree for 2009 Proceed decision (N = 88).

References

Admissions Policy. Undergraduate admission into the University of KwaZulu-Natal in 2009 and there after (2008, August 26). Retrieved July 14, 2008, from University of Kwazulu-Natal University Web site: http://registrar.ukzn.ac.za/Uploads/826711fb-1bb3-4771-a27f-d092fd9fa74b/UKZN%20Admissions%20Policy%202009%20&%20thereafter.pdf

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

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. G. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth.

Centre for Science Access (2005, November). *School Plan for the Centre for Science Access: Progress report.* Unpublished report, Centre for Science Access, University of KwaZulu-Natal.

Department of Education (2006). *Amended norms and standards for school funding. (Notice No.869* of 2006). (Government Gazette 29179) Pretoria: Department of Education. Retrieved February 20, 2010, from www.acts.co.za/south_african_schools_act_1996.htm

Department of Education. (2006b). *Funding for foundational provision in formally approved programmes: 2007/8 to 2009/10*. Document forwarded from the DOE to the Vice-Chancellor of UKZN, and subsequently to faculty staff on 23 May, 2006.

De'ath, G., & Fabricius, K. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology, 81*(11), 3178-3192.

Downs, C. (2005). Is a year-long access course into university helping previously disadvantaged black students in biology? *South African Journal of Higher Education, 19*(4), 666-683.

Downs, C. (2010). Increasing equity and compensating historically academically disadvantaged students at a tertiary level: Benefits of a Science Foundation Programme as a way of access. *Teaching in Higher Education*, 15(1), 97-107.

El-Emam, K., Goldenson, D., McCurley, J., & Herbsleb, J. (2001). Modelling the likelihood of software process improvement: An exploratory study. *Empirical Software Engineering, 6*, 207-229.

Elmitiny, N., Yan, X., Radwan, E., Russo, C., & Nashar, D. (2010). Classification analysis of driver's stop/go decision and red-light running violation. *Accident Analysis and Prevention, 42*(1), 101-111.

Grayson, D.J. (1993). *Summary and overview of the programme*. *Unpublished report,* Science Foundation Programme, University of Natal.

Grayson, D. (1996). A holistic approach to preparing disadvantaged students to succeed in tertiary science studies. Part 1. Design of the Science Foundation Programme (SFP). *International Journal of Science Education, 18*(8), 993-1013.

Grayson, P. (1997). Academic achievement of first-generation students in a Canadian university. *Research in Higher Education*, *38*(6), 659-676.

Grussendorff, S., Liebenberg, M., & Houston, J. (2004). Selection for the Science Foundation Programme (University of Natal): The development of a selection instrument. *South African Journal of Higher Education, 18*(1), 265-272.

Hayden, S., Hayden, M., & Gamst, A. (2005). What characteristics of applicants to emergency medicine residency programs predict future success as an emergency medicine resident? *Academic Emergency Medicine, 12*(3), 206-210.

Kirby, N. (in prep a). Performance of Science Access students in mainstream revisited. Concerns and considerations.

Kirby, N. (in prep b). Foundation Biology and selection model mismatch: Implications for curriculum development.

Kloot, B., Case, J., & Marshall, D. (2008). A critical review of the educational philosophies underpinning Science and Engineering foundation programmes. *South African Journal of Higher Education, 22*(4), 799-816.

Lourens, A., & Smit, I. P. (2003). Retention: Predicting first-year success. *South African Journal of Higher Education, 17*(2), 169-176.

Ma, X. (2005). Growth in mathematics achievement: Analysis with classification and regression trees. *The Journal of Education Research, 99*(2), 78-86.

Mabila, T., Malatje, S., Addo-Bediako, A., Kazeni, M., & Mathabatha, S. (2006). The role of foundation programmes in science education: The UNIFY programme at the University of Limpopo, South Africa. *International Journal of Educational Development, 26*, 295-304.

Morris, C., & Fynn, R. (2003, 26th July - 1st August). *Regression tree models for management of tall grassveld, South Africa.* Paper presented at the VIIth International Rangelands Congress, Durban, South Africa.

Parkinson, J. (2000). Widening access to tertiary science study: the 'augmented' model. *South African Journal of Science, 96*(May), 213-216.

Parkinson, J. (2005, March). Report on language testing. Unpublished Report, Centre for Science Access, University of KwaZulu-Natal.

Pascarella, E., & Terenzini, P. (1998). Studying college students in the 21st century: Meeting new challenges. *The Review of Higher Education, 21*(2), 151-165.

Rollnick, M. (2006). *Successes in science access internationally: A survey of programmes.* In E. Gaigher, L.Goosen & R. de Villiers (Eds.), Proceedings of the14th Annual SAARMSTE Conference (pp. 613-619). Pretoria: University of Pretoria, South Africa

Southway-Ajulu, F. (2005). Report on the exercise undertaken to analyse student cohort data on the Science Foundation Programme (Pmb and Westville) and the extended programme. Unpublished report, Quality Promotions Unit, University of KwaZulu-Natal.

SPSS Classification Trees TM 13.0. User's Guide (2004). Chicago: SPSS (Inc). Retrieved 2009, January 10, from https://softwarerep.ukzn.ac.za/

SPSS Base (Version15) (2006). Chicago, IL: SPSS (Inc).

Trowler, P. (2008). *Cultures and change in higher education: Theories and practices*. London: Palgrave Macmillan.

Umalusi (2008). *Directives for certification National Senior Certificate (Schools).* Pretoria:Umalusi. Retrieved February 2, 2009, from http://www.umalusi.org.za/ur/publications/20081204directivesforcertificationNSC.pdf

Umalusi (2010). *Normalization of 2008 grades and symbols, and 2009 levels*. Presentation of 2008 Maintaining Standards project. Pretoria: Umalusi

van der Flier, H., Thijs, G., & Zaaiman, H. (2003). Selecting students for a South African mathematics and science foundation programme: The effectiveness and fairness of school-leaving examinations and aptitude tests. *International Journal of Educational Development, 23*, 399-409.

Wood, L. & Lithauer, P. (2005). The 'added value' of a foundation programme. *South African Journal of Higher Education*. *19*(5), 1002-1019.

Zaaiman, H. (1998). *Selecting students for mathematics and science: The challenges facing higher education in South Africa*. Pretoria: Human Sciences Research Council.

Appendix A

Values of Senior Certificate symbols and calculation of the Admission Points Score (APS) before the implementation of the NSC in 2008.

Symbol	Percentage achieved in school subject*	Points value for calculation of UKZN Admission Points Score (APS)**	
		Higher Grade	Standard Grade
A	80% - 100%	8	6
B	70% - 79%	7	5
C	60% - 69%	6	4
D	50% - 59%	5	3
E	40% - 49%	4	2
F	30% - 39%	3	1
G	<30%	-	-

* Umalusi (2005)

** UKZN (2009).

For students to have been awarded a Senior Certificate with matriculation endorsement they need to have taken a minimum of 6 subjects at Higher or Standard Grade; two of these were required to be official languages, both requiring a pass (above F) at Higher Grade. At least two of the remaining subjects should have been passed at Higher Grade; a minimum aggregate requirement is also applied (Umalusi, 2005).
All 6 subject results are included in the calculation of the Admission Points Score (APS). If a 7th subject was passed with a symbol of at least 'E' on HG or 'D' on SG, a bonus of 2 points is added to the point score (UKZN, 2009).

Note. Students scoring below F are only considered for the Access foundation streams under special circumstances (see minimum criteria for selection in Table 1 above).

Appendix B

Values of National Senior Certificate (NSC) levels and calculation of the Admission Points Score (APS)

NSC Rating (level of performance)	Description of competence	NSC percentage*	Calculation of Admission Points Score (APS)**	
			Points value	Adjusted percentages for UKZN
			8	90% - 100%
7	Outstanding	80% - 100%	7	80% - 89%
6	Meritorious	70% - 79%	6	70% - 79%
5	Substantial	60% - 69%	5	60% - 69%
4	Adequate	50% - 59%	4	50% - 59%
3	Moderate	40% - 49%	3	40% - 49%
2	Elementary	30% - 39%	2	30% - 39%
1	Not achieved	<30%	1	<30%

* Umalusi (2008)

** UKZN (2009).

For students to have been awarded a National Senior Certificate (NSC) that fulfils the minimum requirements for admission to a bachelor's degree (NSC Deg), they need to have taken a minimum of 7 subjects (including Life Orientation). A minimum of level 2 must have been achieved in the language of learning and teaching of the higher education institution (English in the instance of UKZN). In addition, an achievement rating of 4 or better in four subjects from a designated list (Umalusi, 2008) is required.
Life Orientation is not included in the calculation of the Admission Points Score (APS) at UKZN, although a minimum of a level 4 is required for this subject. The remaining 6 subject results are included in the calculation of the APS. UKZN recognizes academic excellence by awarding 8 points to a subject with a performance level of 90-100% (UKZN, 2009).

Appendix C

Normalization of values of Senior Certificate symbols and NSC levels for the calculation of the Admission Points Score (APS) (Umalusi, 2010).

Higher Grade (400) National Senior Certificate level system Standard Grade (SG)