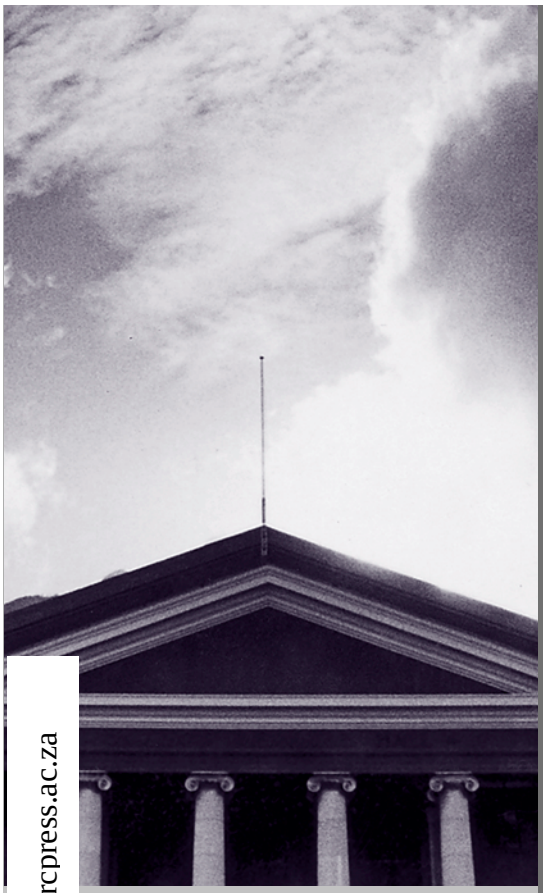
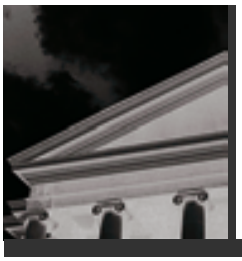


Part One
Background to the Study

1. Origins and rationale of the study
2. Review of the literature
3. The research problem
4. The research design



1. Origins and Rationale of the Study

The proposal for the first South African Student Choice Behaviour project arose out of a seminar in March 2001 under the aegis of the European and South African Group established to support higher education (HE) in South Africa. This Group comprises representatives of the South African national Department of Education and a consortium of European researchers from:

- The Netherlands – the Centre for Higher Education Policy Studies (CHEPS) and the Faculty of Education at the University of Twente;
- Norway – the Norwegian Institute for Studies in Research and Higher Education (NIFU)[1] and the Norwegian Ministry of Church Affairs, Education and Research (KUF); and
- Finland – the University of Turku / RUSE.

The Group's support for HE has been realised through the launching of two projects: Leadership in Higher Education, the aim of which is to improve leadership in South African HE; and Student Choice Behaviour, the aim of which is to investigate the factors affecting student choice with regard to HE with a view to influencing policy formulation in the areas of, among others, career counselling, increasing matriculation outputs, teaching and learning improvement, admission to HE institutions, and HE programme planning.

1.1 Purpose of the study

This report focuses on the outcomes of Phase One of the Student Choice Behaviour project conducted by the Research Programme on Human Resources Development (HRD) at the Human Sciences Research Council (HSRC) in August 2001. Commissioned by the Department of Education (DoE), the study seeks to understand the factors affecting the choices Grade 12 learners make with regard to entering higher education (HE), institution type (university or technikon) and specific institution, and programme of study. The project is the first in a planned series of studies which in its entirety investigates the transition from school to work.

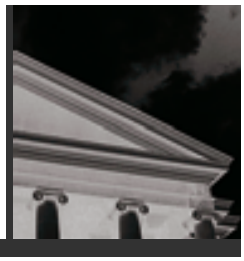
This first study provides an initial indication of learner choices, an understanding of which will facilitate planning at secondary education level in the areas of programme (i.e. subject) focus, career guidance, learning improvement and teacher upgrading – in particular, enabling Department of Education planners and school management to:

- Devise strategies to increase the number of Grade 12 learners entering HE (at present there is a 15% participation rate[2] in public HE in South Africa) in subjects that will lead to HE qualifications that address the national demand for specific skills;
- Plan more effective career guidance programmes at the Grade 11 and 12 levels; and
- Target specific subjects and subject areas for teacher improvement programmes.

In Phase Two of the project, which will take place in the HSRC 2002/3 financial year (April 2002 to March 2003), those students enrolled for HE study who participated in Phase One of the project will be tracked into HE institutions. A survey will seek to confirm their choices of institution and study programme and gauge their perceptions about their career plans. Phase Three will see the launching of a study investigating the factors affecting the subject choices of Grade 9 learners for their upper secondary

1 The assistance of Per Olaf Aamodt and Vibeke Opheim of NIFU in the conceptualisation and instrument design phases of the project is gratefully acknowledged.

2 The participation rate is calculated, following the United Nations Education, Scientific and Cultural Organization (UNESCO) definition, as the number of students aged 20 to 24 enrolled in the HE system (DoE, 2001a). While the *National Plan for Higher Education* (DoE, 2001a) figure is 15%, a recalculation in late 2001 to include technical college students registered for HE programmes (N4 – N6) boosts the rate to about 18% (Bunting, 2002).



education. A triangulation of the results of the three components of the project – the Grade 12 learner choice survey, the first-year student choice survey, and the Grade 9 learner choice survey – will provide a richly textured understanding of the factors affecting learner choice with regard to HE.

1.2 Justification for the study

The immediate impetus for the present study comes from the South African *National Plan for Higher Education* (DoE, 2001b), in particular two of its objectives:

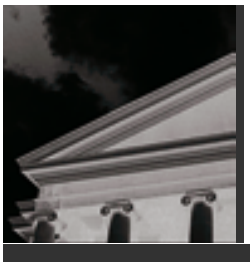
- To increase the participation rate in HE from 15% to 20% in the long term (10–15 years) in an attempt to address both the imperative for equity as well as changing human resource and labour needs; and
- To shift the balance in enrolments in the short to medium term (5–10 years) between the humanities, business & commerce, and science, engineering & technology (SET) from the current ratio of 49%:26%:25% to 40%:30%:30% respectively.

The first objective should be understood in the context of National Commission on Higher Education (NCHE) predictions that the participation rate in South African HE would increase from 20% in 1996 to 30% in 2005 (NCHE, 1996). As the *National Plan* points out, however, the figure of 20% was erroneous as it was based upon the 1991 Census, which was flawed because it excluded "homelands" under the apartheid regime; a more accurate figure (for 1996) is 17%. Not only were the NCHE estimates inflated because of an erroneous reference point, however; but enrolments, though they increased by 135 000 headcounts between 1993 and 1998, began levelling off in the last two years of this period (growing by only 3%) and then declined by 4% between 1998 and 2000. The decrease in enrolments is attributable largely to two factors: the sharp decline in throughputs from the schooling system (a 23% decline from 1994 to 2000), particularly of learners with matriculation exemption (a pre-condition for entry into universities and even in some case to technikons); and a sizeable fall in the retention rate in HE (i.e. the proportion of students in a given year who re-register in the following academic year). An average of 20% of undergraduate and postgraduate students drop out of HE each year (DoE, 2001b).

The second objective should be read against the backdrop of enrolment patterns in business & commerce and SET in relation to the humanities. The period 1993–1999 saw a shift in enrolments from the humanities to business & commerce; in the same period, however, enrolments in SET remained fairly constant. More specifically, the enrolment proportions showed a decline from 57% to 49% in the humanities, a rise from 19% to 26% in business & commerce, and a very slight rise from 24% to 25% in SET. A major constraint on increasing enrolments in business & commerce and in SET is the paucity of matriculants with a higher grade pass in mathematics – in 2000, only 19 327 (or 7% of the total number of matriculants for that year).[3]

An understanding of the background particularly to the second of the *National Plan* objectives mentioned above is critical to the policy formulation intention of the Student Choice Behaviour study. If the factors that lead learners to choose study programmes can be understood, policies and concomitant strategies can be devised to influence students in selecting study programmes that will result in the achievement of qualifications that address the national demand for an appropriate balance of skills in the workplace.

3 This situation did not improve markedly in 2001: there were 19 504 higher grade (HG) Mathematics passes – still about 7% of all matric candidates. However, the Mathematics pass rate improved slightly: 55.9% of all those who wrote HG Mathematics passed on the higher grade, in comparison with 50.2% in 2000 (DoE, 2002).



2. Review of the Literature

2.1 Branching points in learner decision-making

In an ideal world, student choice behaviour would be a matter of unfettered choice. In reality, however, students make choices on the basis of the options open to them – options constrained both by personal and socio-economic circumstances (race, gender, socio-economic background, etc.) and institutional factors (institutional selection mechanisms, institutional capacity, etc.).

In the South African secondary education system, there are two critical points at which students make decisions. The first is in the course of the Grade 9 year,[1] when students decide on the combination of subjects they will take for the remainder of their secondary schooling; and the second is in the Grade 11 or Grade 12 year, when students decide whether to enter higher education, if so, whether to enter a technikon or a university, and what course of study to pursue. While decision-making is a fairly workable notion for Grade 11 or 12 learners, however, Grade 9 learners are in many instances constrained by their own ignorance of the consequences of selecting subjects for their upper secondary education,[2] or indeed by their academic performance in the subjects they have taken to date. Nevertheless, these branching points, as Boudon (1974) calls them, constrain students in different ways. For example, for young students (notionally aged 15) from low socio-economic backgrounds there are costs involved not only in choosing (in Grade 11 or 12) to enter higher education institutions – where the costs have to do with alienation from one's roots – but even in deciding whether to continue with upper secondary education (or Further Education and Training [FET], in the language of the National Qualifications Framework), and if so what combination of subjects to pursue. The choice of subjects for FET is made at the Grade 9 level; and since this choice of subjects may well determine whether a student enters higher education after the Grade 12 year, no study on student choice behaviour can afford to ignore this critical branching point.

The decision to enter higher education may be regarded as a multi-stage process involving a series of successive decisions finally resulting in enrolment in a higher education programme (Hossler *et al.*, 1989). Generally, three broad stages can be distinguished in the process:

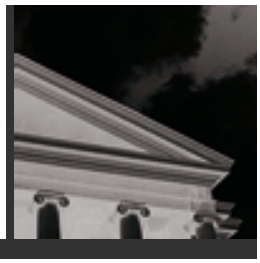
- Deciding to enter higher education;
- Selecting a particular institution and programme of study; and
- Persisting in higher education (Campaigne & Hossler, 1998).

The first broad stage is itself divisible into two: the decision to enter those upper secondary programmes leading to higher education (making the choice in Grade 9 about which subjects or cluster of subjects to pursue until completion of the matriculation certificate); and the decision (taken either in Grade 11 or Grade 12) to enter higher education (or not) after completing upper secondary education.

2.2 Student choice models

Three types of models have been developed to explain student choice of institution for higher education study: economic models; sociological models; and information processing models. These are reviewed below.[3]

1 Notionally at the age of 15 – though frequently, because of commencement of schooling at an advanced age, interruption of schooling, or repetition of previous years of study, at age 16 or 17.
2 The fact that, in 2001, 4 350 candidates passed matric with merit but without endorsement, thereby limiting their options with regard to entry into HE, may be attributable to poor subject choice guidance early on in their high school careers.
3 Section 2.2 builds upon a review of the literature on student choice behaviour by Aamodt (2001).



2.2.1 Economic models

Economic models are based on the notion that a student maximizes utility using some kind of cost-benefit analysis. Costs include both direct costs, such as tuition, books, opportunities, and cost of living, and indirect costs, such as the cost of leaving home and friends. Such models focus on how individuals differ, by virtue of individual characteristics, in terms of those variables that are important in the choice of a higher education institution.

While the college-going models of such authors as Bishop (1977) and Manski & Wise (1983) attempt to explain choice between higher education and other alternatives, the choice of a particular institution and programme of study is addressed by such authors as Kohn *et al.* (1976) and Chapman (1984). The most important factors identified in these models that explain student choice are tuition, net tuition (tuition minus financial support), subsistence costs, foregone earnings, expected future earnings, family background, gender, average ability, and aspirations of neighbourhood peers.

2.2.2 Sociological models

Sociological models focus on the identification and interaction of variables as students make decisions about entering higher education. According to this model, there are a number of factors that influence student enrolment decisions, including socio-economic status (especially family background), academic ability, significant others, secondary school performance, educational aspirations, motivation to succeed, and secondary school characteristics.

2.2.3 Information processing models

The most powerful models for explaining student choice behaviour combine the indicators identified in the economic and sociological models to provide a conceptual framework that attempts to account for the effects of policy-making interventions. Hossler *et al.* (1999) illustrate that information processing models attempt to describe the variety of economic and social forces that affect individual student decision making *in order to identify appropriate intervention strategies that will influence student choice.* They identify four major models of student choice, outlined below.

2.2.3.1 The Jackson model (Jackson, 1982)

There are three stages in this model. In the first, the *preference* stage, academic achievement is shown to correlate strongly with educational aspiration. In the second, the *exclusion* stage, economic factors (location, costs, and academic quality) are used to exclude institutions. And in the third, the *evaluation* stage, the remaining institutions are assessed on the basis of their qualities.

2.2.3.2 The Chapman model (Chapman, 1984)

While the Jackson model is student-oriented, Chapman suggests that *student* characteristics (socio-economic status, scholastic aptitude, educational aspirations, and academic performance) interact with external influences (significant others [friends,

parents, secondary school personnel], institutional characteristics, and institutional marketing) to create a student's general impression of higher education life.

2.2.3.3 The Hanson & Litten model (Hanson & Litten, 1982)

This model characterizes student choice as a continuing process wherein there are five key phases: *aspiration; commencement of the search process; information gathering; submission of applications; and enrolment.* The variables affecting the choice process are background characteristics (parental income, parental education, race, and gender), personal characteristics (academic ability, class rank, and self-image), secondary school characteristics (social composition, programmes offered, curriculum), and higher education institution characteristics (costs, size, programmes offered, and timeousness in responding to queries).

2.2.3.4 The Hossler & Gallagher Model (Hossler & Gallagher, 1987)

This model is based upon a three-stage conception that locates the student at the centre: the *predisposition* stage (decision to enter higher education); the *search* stage (learning about and comparing institutions); and the *choice* stage (completing applications and choosing an institution). The model uses the term *information processing* to indicate the self-reflective, holistic nature of the process of choice, which is located within a social setting dubbed the social capital of students. Social capital includes not only student background characteristics (for example, ability) and family background characteristics (for example, parental income), but also the preferences and attitudes transmitted to children and the ways in which parents shape their children's futures (through, for example, reading to them, modelling reading habits, and encouraging critical thinking). The social capital concept is premised upon the interaction between the three major socializing agents of students: the family; the peer group; and the school. In this context, higher education choice is conceived of as a process requiring the interplay of different variables at different times – culminating in the final choice of institution and programme of study.

2.2.4 Key findings of information processing models

The key findings of studies conducted by Jackson (1982), Chapman (1984), Hanson & Litten (1982), and Hossler & Gallagher (1987) are the following:

- Family background has a moderate to strong effect on learner choice at all three stages (preference, exclusion, and evaluation) (Jackson, 1982);
- Location of HE institution, cost of HE, access to information about HE and job prospects after graduation have a moderate to strong effect on learner choice at the exclusion and evaluation stages (Jackson, 1982);
- Academic achievement has the strongest correlation with educational aspiration – that is, the preference (to enter HE) stage – and a moderate to strong effect on learner choice at the exclusion and evaluation stages (Jackson, 1982);
- Student characteristics – socio-economic status (SES), scholastic aptitude, educational aspirations and academic performance – and external influences – significant others (friends, parents, teachers), HE institution characteristics (cost, location, programmes)

and HE institution marketing efforts (written information, campus visits, recruiting) interact to form a learner's general impression of HE life (Chapman, 1984);
- Family income has a strong effect on learner choice at the presearch stage (Chapman, 1984);
- Background characteristics – parental education, parental income, population group, and gender – personal characteristics – social composition and school subjects – and HE institution characteristics – cost of HE, size of institution, programmes offered, response time to student queries, and financial aid – together affect learner choice (Hanson & Litten, 1982);
- Social conditions strongly influence learners during the search stage, while academic and social attributes of HE institutions have a strong effect during the choice stage (Hossler & Gallagher, 1987); and
- Background characteristics are strongly correlated with the predisposition (to enter HE) stage (Hossler & Gallagher, 1987).

All these studies were conducted on the North American continent, however – which has implications for attempts at transplanting the findings into the South African context.

2.2.5 Financial incentives

The literature on the influence of financial incentives on student choice is divided. While some studies (for example, Campaigne & Hossler, 1998; Kane, 1995; McPherson & Schapiro, 1998) demonstrate a correlation between increase in fees and decrease in enrolment of (particularly lower socio-economic status) students, on balance the majority of studies (reviewed by Leslie & Brinkman, 1988 and Heller, 1997) demonstrate that price increases lead to only minor decreases in enrolment numbers, with grants leading to only slight increases in student demand. In short, tuition fees and student support have only a limited effect on student choice.

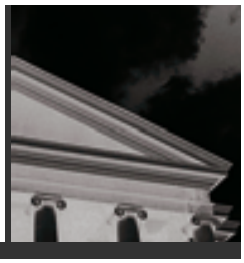
What is apparent, however, is that students from different socio-economic backgrounds differ in their responses to price incentives. The only other study (besides the present study) of student choice to be conducted in the South African context (Steenekamp, 2000) – a survey of the factors affecting Grade 11 and Grade 12 learner choices with regard to HE amongst 2 022 learners in 14 schools in the Eastern Cape – confirms that nearly half of African and coloured learners (49%), most of whom (though the study does not investigate this) would probably come from low socio-economic backgrounds, cite lack of access to finance as their main reason for not intending to enter HE, while a high percentage (44%) of white learners, most of whom would come from middle to high socio-economic backgrounds (again, the study does not confirm this), cite wanting to travel abroad before deciding whether to enter HE as their main reason for not proceeding to HE.

How financial status affects learner choice needs to be addressed not only in a provincial context, however, but in the *South African* context, in a way that allows for cross-tabulation not only of population group but of socio-economic status (SES) with financial constraint as a disincentive to entering HE.

2.2.6 Relevance of the literature to the present study

The project reported on here, as intimated above, constitutes the first *national* study of student choice behaviour in South Africa. In one sense, the literature reviewed above (the Steenekamp report aside), which in describing the state of the art pulls together the major *trends* in research on student choice, has a limited usefulness for an *initial*, as yet unbenchmarked, study conducted in an educational environment in which the examination of the effects of a combination of economic, sociological and information-processing factors on learner choice is a new construct. In another sense, however, the efficacy of a research model evolved largely in a developed world context needs increasingly to be tested in a developing country context if a global view of student choice behaviour is to be obtained.

What is abundantly evident from the body of research on the transition from secondary education to HE reviewed above is the shift from a focus on educational outcomes as the key predictor of entry into HE to an understanding of the interrelatedness of a number of variables in the decision-making process. The observation of Gayle *et al.* (2000: 62) that 'The effects of these (individual and social background) factors are (best) understood as part of an ongoing social process that underpins the young person's educational career' hints at the need for longitudinal studies into the influences of different social factors on student choice. In this context, the present study provides baseline data and a sizeable student cohort for subsequent national studies into the effect of a variety of factors on learners' choices with regard to entry into HE, the pursuit of postgraduate study, and entry into the labour market.



3. The Research Problem

3.1 Research question and hypotheses

3.1.1 Research question

The central research question to be addressed in the study is: *What factors affect the choices of Grade 12 learners in schools across South Africa with regard to HE?* Embedded within this question are the following subsidiary questions:

- What factors affect the choices of Grade 12 learners with regard to entry into HE?
- What factors affect the choices of Grade 12 learners with regard to institution type and specific institution?
- What factors affect the choices of Grade 12 learners with regard to programme of study?

The project investigates the transition from school to post-school *with particular reference to entry into HE*: while the factors affecting the choices of learners *not* to enter HE clearly throw light upon responses to the question 'What factors affect your choice with regard to entry into HE?', the study is not primarily concerned with those learners not intending to enter HE. A separate, related, study of those factors affecting such learners, however, would enrich an understanding of HE entry factors, since the two sets of factors exist in an asymmetrical relationship to each other. (Such a study to complement Phase One of the Student Choice Behaviour project is planned for the 2002/3 financial year.)

3.1.2 Research hypotheses

3.1.2.1 Background to statement of hypotheses

There are two kinds of factors that affect the choices of Grade 12 learners with regard to HE as posited by the questionnaire (reproduced in the Appendix): attitudinal factors – which are subjective and future-oriented; and *status quo* or 'given' factors – which are objective and largely biographical. For ease of analysis, the positivist terms *subjective* and *objective* are used to differentiate these two kinds of factors.

Table 3.1 groups specific factors in the questionnaire into groups of *generic* factors, which in turn are grouped into larger categories of influence through a conceptualisation of the learner's process of choice with regard to HE in a spatio-temporal way, as follows:

Individual ⟶ Home ⟶ School ⟶ HE institution ⟶ Work ⟶ Country

The learner's choice process, according to this conception, is located within a widening gyre of spheres of influence, at the centre of which is the learner as individual (a psycho-social construct shaped by a combination of genetic and social factors) located within the home (where the influence of parents and siblings is strongest) interacting with:

- The school – where peer pressure, academic performance, teachers, and career guidance are shaping influences;
- HE institutions – where both institutional factors like location, reputation, modes of study and quality of provision as well as study programme factors like availability of places, theoretical versus practical orientation, and subject appeal are important influences;
- work – notions of work in relation to HE and as an alternative to HE, as well as work situation ten years hence; and

- the country – where the 'state of the nation' and perceptions of HE in relation to perceptions about the country as a whole exert an influence on learner choice.

Table 3.1: Matrix of factors affecting learner choice with regard to HE

GYRE OF INFLUENCE	GENERIC INFLUENCE	SPECIFIC INFLUENCE	
		Objective	Subjective
Individual	Personality type		Indecision / lack of direction Need for independence Insecurity with regard to ability
	Gender	Gender	
	Physical disability	Physical disability	
	Population group	Population group	
Home	Parents	Employment situation Education	Influence on learner
	Finance	Parental income	Access to finance
		Home province	Home province
	Siblings	Studying Graduates	
	Family situation		Family responsibility
School	Academic performance	Academic performance (Grade 11)	Academic performance
	School	Provision of career guidance	Provision of career guidance Teacher influence
	Peer pressure		Peer pressure (classmate + friend) Romantic attachment
Higher Education	HE institution		Quality of HE provision Location of HE institution Reputation of HE institution / programme HE institution a launch pad for study abroad Personal security of learner at HE institution Sporting facilities at HE institution Opportunity to study via correspondence

Table 3.1: Matrix of factors affecting learner choice with regard to HE (cont.)

GYRE OF INFLUENCE	GENERIC INFLUENCE	SPECIFIC INFLUENCE	
		Objective	Subjective
Higher Education	HE study programme		Intrinsic interest in a field of study Availability of places in HE institution / programme Practical orientation of field of study Theoretical orientation of field of study Academic interest in matric subjects
	Information about HE		Access to information about HE
Work	Work		Desire to work HE increasing chances of employability HE leading to higher income Job offer Contribution of study programme to national development
	Work situation ten years hence		Large income Status Use of skills Alignment with interests Independence Teamwork Job security Usefulness to society Philanthropy Flexibility
Country	State of the nation		Crime situation Poverty Quality public education Quality public health care HIV/AIDS under control

3.1.2.2 Hypotheses of the study

A synthesis of the research on student choice behaviour conducted by Jackson (1982), Chapman (1984), Hanson & Litten (1982), and Hossler & Gallagher (1987) reveals that:

- Background characteristics exert a strong influence on learner choice at every stage of the process, but particularly at the initial stage (presearch, predisposition, or preference); and

- Learner choice is a *process* subject to the influence of a number of variables *in combination*, and it is therefore difficult to isolate one specific factor which exerts the strongest influence on learner choice.

Given the legacy of *apartheid*, the effects of which the majority of South Africans still feel on a daily basis, however, there are certain factors in the South African context which may well prove to have an overriding effect on learner choice with regard to HE. The first of these is population group, the disaggregation of which under *apartheid* spelled fourth-class citizenship for (black) Africans, third-class citizenship for coloured people, second-class citizenship for Indians; and the second, closely correlated with population group, is socio-economic status (SES). An examination of current unemployment rates and income distribution in South Africa reveals that the country has:

- Significant unemployment (an official unemployment rate of 26.4% – partially disaggregated to 31.1% for Africans and 6.6% for whites)[1];
- Vast disparities in income distribution by population group (32.4% of African workers[2] earn R1-R500 per month and 1.2% earn R8 001 or more per month, while amongst white workers these percentages are almost reversed – 2.4% of white workers earning R1-R500 per month and almost a quarter (24.1%) earning R8 001 or more per month); and, concomitantly,
- One of the highest Gini indices[3] in the world (in 1993/4 the figure was 59.3; the index is higher only in Guatemala – 59.6, Brazil – 60.0, Nicaragua – 60.3, Swaziland – 60.9, Central African Republic – 61.3, and Sierra Leone – 62.9).[4]

Against this backdrop, one of the specific hypotheses of the study must be that financial status, correlated with population group, is a significant predictor of predisposition towards entry into HE.

In the light of the literature reviewed in Chapter 2 and the matrix of factors affecting learner choice with regard to HE developed in Table 3.1, we are able to formulate two hypotheses of the study. They are that:

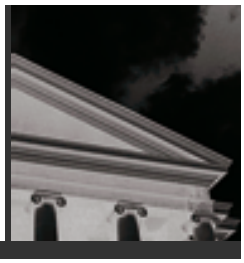
- A range of factors exerts an influence on student choice behaviour with regard to HE; and
- Family background, in particular SES, is strongly correlated with learner choice particularly at the predisposition (to enter HE) stage.

1 All statistics in these first two bullet points are derived from the *Labour Force Survey* of February 2001 (StatsSA, 2001a&b).

2 Workers, categorized into employers, employees, and those in self-employment, are defined as those persons who 'performed work (at least one hour per week) for pay, profit or family gain in the seven days prior to the household survey interview, or who were absent from work during these seven days, but did have some form of paid work during this time' (Stats SA, 2001a: xii).

3 The Gini index, used to measure equality and inequality within countries or between groups of people, has a value between zero (perfect equality) and 100 (perfect inequality).

4 These Gini indices reflect a combination of the World Bank and United Nations Development Programme (UNDP) reports for 2001. Because different indices were measured in different years, however (countries do not necessarily update these figures annually), these comparisons should serve as a guide only.



4. The Research Design

The research for Phase One of the Student Choice Behaviour project involved the collection of primary data by means of a cross-sectional quantitative survey research design. The survey was conducted in a sample of schools within all nine provinces during August 2001. A quantitative survey research design was used primarily for two reasons: firstly, it allowed the collection of data on factors influencing the behaviour of a large target population; and second, it allowed the findings to be generalised to the entire population of Grade 12 learners in the country. The key methodological steps for the survey design were the development of the instrument, sampling, and fieldwork.

4.1 Research methodology

4.1.1 Questionnaire design

The survey was administered through the deployment of a self-completion questionnaire, which was designed to be completed by Grade 12 learners in a group situation under the supervision and with the guidance of an HSRC-trained fieldworker.

The multi-column layout deployed in the questionnaire allows a concentration of 583 variables to be captured in the most economical way possible, while the predominance of closed-ended items (only 9 of the 65 questions are open-ended) facilitates ease of completion (and of course data capture) of the questionnaire. The 65 questions in the questionnaire are distributed across the following sections:

- Present school situation;
- Preferred and anticipated life situations one year hence;
- Intention to study at a HE institution;
- Application to study at a HE institution;
- Choice of HE institution;
- Choice of study programme;
- Anticipated life (particularly work) situation ten years hence;
- Biographical information; and
- Unfettered choice with regard to HE and work.

The main features of the logic underpinning the questionnaire design are the following:

4.1.1.1 Introduction to the questionnaire

Two tables on the first page of the questionnaire allow for the capture of learner details. The first – completed by the fieldworker – captures the unique six-digit code assigned to each respondent. This learner code comprises three elements: the province in which the respondent's school is located (provinces are coded from 01 to 09 in alphabetical order); the school at which the respondent is a Grade 12 learner (schools are coded from 001 to 300, in alphabetical order within provinces); and a two-digit code for each of the learners (up to a maximum of 48) at each school participating in the survey. The second table – completed by the learner – captures the personal details of each learner: full names; surname; identity number; date of birth; and name of school. The remainder of the first page outlines the purpose of the survey and instructions for the completion of the questionnaire.

4.1.1.2 Section progression

The questionnaire itself is divided into nine sections, moving from the respondent's present school situation (Section 1) to the learner's envisaged situation one year hence (Section 2), to the learner's intention to study at an institution of higher learning (Section 3), to the learner's actual or intended application to an institution (Section 4), to the learner's choice of institution for study (Section 5), to the learner's choice of study programme (Section 6), to the learner's perceptions about work and his / her envisaged life situation ten years' hence (Section 7), to the learner's personal situation at present (Section 8), to the learner's unfettered choices with regard to higher education and work (Section 9).

4.1.1.3 Rationale behind section progression

Because there are some questions in the 'Personal Information' section to which some learners may respond more sensitively than others (for example, questions about population group, parental employment and income, and living conditions), this entire section is relegated to the end of the questionnaire, constituting the penultimate section. The sequencing of the nine sections of the questionnaire proceeds as follows:

- The first section deals with that which is of most immediate academic concern to learners, their present school situation – in particular, the subjects in which they will be writing their matriculation (matric) examinations.
- Section 2 extrapolates the learner's situation from school to post-school one year hence through a comparison of preferred and envisaged life situations, introducing into these two scenario conceptions the notion of HE.
- Section 3 foregrounds the notion of study at a HE institution adumbrated in Section 2 by asking the respondent to commit him / herself at the outset through response to a yes-no question (3.1) to intention to enter HE within the next three years – the remainder of the section probing the factors that have influenced the learner's decision to enter HE or not to enter HE, or the learner's indecisiveness on the issue.
- The filter at the beginning of Section 4 directs only those respondents intending to enter HE to answer Sections 4, 5 and 6 – thereby filtering out all Section 3 respondents either unsure about entering HE or not intending to enter HE and directing them to Sections 7 to 9.
- While Section 4 focuses on the application process, Sections 5 and 6 home in on the choice of institution and study programme respectively. There is a progressive narrowing of focus, then, in Sections 4, 5 and 6 towards the actual anticipated field of study of the prospective HE learner.
- Section 7 abstracts the respondent from a concern with HE (whether in the notional sense broached in Section 3 or the very particular sense plumbed in Section 6) to his / her future career prospects and life situation ten years' hence – thereby implying (in the logic of the questionnaire) that HE could occupy the intervening nine years of the learner's life between his / her situation one year hence (Section 2) and ten years' hence (Section 7). An implicit comparison of these temporally disjoined life situations is hereby invited.
- Finally, Section 9 abstracts the respondent even further from his / her present and future life situations by positing the notion of choice of HE and employment within a constraint-free context, thereby allowing for the cross-tabulation of responses to

the question '(in an ideal world) Would you choose to study at a university or technikon?' (question 9.1) with the responses to 'Are you planning to study at a university or technikon within the next three years?' (question 3.1).

4.1.2 Sampling

The sampling frame used to design the sample was a database containing the names of all secondary schools in South Africa that participated in the 2000 Senior Certificate examinations. A two-stage sampling procedure was deployed: in the first stage, 300 senior secondary schools were selected for the study; and in the second, 48 Grade 12 learners were selected from each participating school.

4.1.2.1 First-stage sampling

Since results were required to be reported at both national and provincial levels, schools had to be drawn in such a way that there was an adequate number from each province in the sample. To ensure that this was the case, the population of 5612 schools was first stratified according to the nine provinces. Following this, the first option was to allocate the sample of 300 schools proportionally to the provinces.[1] This option, however, resulted in too few schools (only five) being allocated to Northern Cape. The second option, consequently, was to apply disproportional allocation, in which Northern Cape schools were over-sampled by 10 (rendering a sample of 15 schools) and the two biggest provinces (in terms of numbers of schools) – KwaZulu-Natal and Limpopo – each under-sampled by 5 schools. The table below indicates the allocation of the sample of 300 schools to the nine provinces.

Table 4.1: Provincial allocation of sample of 300 schools

PROVINCE	SCHOOLS WITH GRADE 12 LEARNERS	PROPORTIONAL ALLOCATION	DISPROPORTIONAL ALLOCATION
Eastern Cape	875	47	47
Free State	324	17	17
Gauteng	598	32	32
KwaZulu-Natal	1 295	69	64
Mpumalanga	389	21	21
Northern Cape	101	5	15
Limpopo	1 310	70	65
North West	367	20	20
Western Cape	353	19	19
Total	**5 612**	**300**	**300**

The number of schools in Limpopo is disproportionately high for a province with a relatively small population: while the percentage of the South African population living in

1 This does not take into account possible different average sizes of schools between provinces or, more specifically, average sizes of the Grade 12 classes. Taking these factors into account would have over-complicated the sampling process, however.

Limpopo is only 12.1% – compared with 15.5% in the Eastern Cape, 18.1% in Gauteng, and 20.7% in KwaZulu-Natal[2] – the number of schools included in the proportional allocation for this province (70) is larger than that for any of the larger-population provinces – which indicates that schools in this province are on average smaller than in other provinces.

Another important factor to take into account was the pass rate in the schools – that is, the percentage of Grade 12 learners who passed the Senior Certificate examination in 2000. The requirement was that the sample of schools be representative in terms of this pass rate. To arrive at such a sample, it was decided to stratify the population of schools into three strata according to pass rate. The first stratum contained all schools with a pass rate of 0% – 33.33%, the second stratum contained all schools with a pass rate of 33.34% – 66.67%, and the third stratum contained all schools with a pass rate of 66.68% – 100%. The numbers of schools allocated to each province were then distributed proportionally across these three strata (within each province) to ensure that these three strata were proportionally represented in the sample of schools. The following table shows the final allocation of the 300 schools after this refinement of Table 4.1.

Table 4.2: Final allocation of sample of 300 schools

PROVINCE	SCHOOLS WITH GR 12 LEARNERS	PASS RATE 0-33.33%	PASS RATE 33.34-66.67%	PASS RATE 66.68-100%	TOTAL
Eastern Cape	875	17	19	11	**47**
Free State	324	5	6	6	**17**
Gauteng	598	6	11	15	**32**
KwaZulu-Natal	1 295	20	24	20	**64**
Mpumalanga	389	5	10	6	**21**
Northern Cape	101	1	5	9	**15**
Limpopo	1 310	16	32	17	**65**
North West	367	3	10	7	**20**
Western Cape	353	1	4	14	**19**
Total	**5 612**	**74**	**121**	**105**	**300**

In order to arrive at a more or less self-weighting sample of learners in each of the 27 explicit strata (9 provinces x 3 pass-rate strata), the allocated number of schools within each explicit stratum was selected with probability proportional to the number of Grade 12 learners who wrote the examination in 2000, and subsequently a fixed number of learners (48) was selected from amongst all Grade 12 learners in the selected schools.

4.1.2.2 Second-stage sampling

After the 300 schools had been selected using the method described above, 48 Grade 12 learners had to be selected in each school. In cases where the total number of Grade 12

2 These figures are taken from the 1996 Census (Stats SA, 1998: 4).

learners was 48 or fewer, all learners had to be included in the sample; in other words, the survey comprised a census of the Grade 12 learners in the school. In cases where the total number of Grade 12 learners exceeded 48, the required number of learners was selected using convenience sampling, the selection of survey participants being left to the discretion of the school principal or a senior member of staff – who was asked to select learners in as random and representative a fashion as possible. This approach was taken for the following reasons:

- Schools in the sample differed according to gender composition: both co-educational and single-sex schools were included;
- Schools differed according to their administrative capacity to select participants according to any pre-determined sampling procedure; and
- The geographical location and / or poor communication infrastructure of certain schools (especially those in inaccessible rural areas) precluded advance notice of the survey being given – which rendered the drawing of a planned learner sample unattainable.

Fieldworker reports on the survey process indicate that the selection process took different forms in different schools. Some schools differentiated by gender (deliberately including equal numbers of boys and girls in the sample), some schools by academic performance (deliberately including learners of different academic abilities in the sample), and some by subject choice (deliberately including learners representing a cross-section of subject areas). Others either deployed a combination of two or more of these stratification strategies, or randomly assembled 48 learners in a classroom on the day of the survey. The project team is confident that the sample was sufficiently large such that the combined effect produced a sample representative of the Grade 12 learner population in the country.

4.1.3 Fieldwork arrangements

4.1.3.1 Access to schools

Official sanction for the survey was sought from the Higher Education Branch of the Department of Education (DoE), which in a letter to the heads of the nine provincial education departments requested their support for the project. These heads subsequently communicated to the district and circuit offices in their provinces the intention of the HSRC to conduct the survey in the schools selected for the study over a two-week period in August 2001. Most schools were apprised of the survey ahead of time, but many in rural areas could not be contacted timeously. In any event, access to schools was readily granted, only a few school principals objecting to the presence of the HSRC, on the grounds of interference with the writing of the preliminary matric examinations, of point-blank refusal to participate in the survey at all (a couple of schools in the Western Cape), or (in one instance only, in Gauteng) of the questionnaire being available in English only (and not in Afrikaans).

4.1.3.2 Recruitment and training of fieldworkers

A total of thirty fieldworkers – one per group of ten schools in each province – was recruited by the project team to facilitate learner completion of questionnaires. Fieldworkers were recruited largely through reference to the cohort of contract staff who

had worked on the Quality Learning Project managed by the Assessment Technology and Education Evaluation Research Programme of the HSRC in 2000. Conveniently, since many of these fieldworkers were provincially based, such recruitment allowed the project to make use of persons familiar, to a greater or lesser extent, with the school districts in the provinces in which they were resident and with the home languages of learners who participated in the study. According to the survey design, each fieldworker conducted the survey in one school per day over a ten-day period.

Fieldworkers were trained according to a cascade model: the Project Manager trained HSRC staff in the HSRC's offices in Pretoria; they in turn trained fieldworkers in the provinces in which they were based. Fieldworker training involved a thorough immersion in the aims and objectives of the project, in the process of completing the questionnaire, and in the process of facilitating the questionnaire completion process by learners – the cumulative intention of which was to maximise the validity of the survey.

4.1.3.3 Response rates

Given the tight timeframe within which the project team had to operate (between the date of the DoE communication with provincial education departments about the project and the date of commencement of fieldwork there was a period of only five weeks), the response rate to the survey was remarkably high. In the period 13 to 24 August 2001 HSRC-trained fieldworkers visited 288 of the 293 schools included in the sample[3] to facilitate the completion of the questionnaire (a response rate of nearly 98%), and of the 14 064 questionnaires taken into the field (48 per each of the 293 schools) we received 12 201 completed questionnaires – a response rate of 86.8%. The figure of 14 064 was based on an estimation of 48 completed questionnaires per school visited; however, since some schools did not have 48 Grade 12 learners, the actual sample size is smaller than 14 064, and the true response rate will therefore be higher than 86.8%. The following table indicates the breakdown of responses by province:

Table 4.3: Summary of school and learner questionnaire response to Grade 12 Learner Choice survey

PROVINCE	NO. OF SCHOOLS RETURNING QUESTIONNAIRES	SCHOOL RESPONSE RATE (%)	NO. OF QUESTIONNAIRES RETURNED[4]	QUESTIONNAIRE RESPONSE RATE (%)
Eastern Cape	44	97.8	1932	89.4
Free State	17	100.0	731	89.6
Gauteng	27	96.4	1219	90.7
KwaZulu-Natal	62	96.9	2636	85.8
Mpumalanga	20	100.0	845	88.0
Northern Cape	14	100.0	557	94.2
Limpopo	64	98.5	2466	79.0
North West	20	100.0	934	97.3
Western Cape	20	100.0	881	84.6
Total	**287**		**12 201**	**86.8**

3 The sample size had to be reduced from 300 to 293 because of incomplete records in the DoE database upon which the sample frame was based – the latitude and longitude correlates for seven schools being unavailable.
4 These figures do not represent the numbers of valid responses, which are reported in the 'Research findings' section below.

Five schools – two in KwaZulu-Natal, one in Gauteng, one in Mpumalanga, and one in Limpopo – were not surveyed, for reasons either of plain refusal or of learners physically not being present at school on the day of the survey, or for logistical reasons (for example, the fieldworker not being able to find or physically get to the school). As a last resort – in cases where no other solution could be proposed – certain schools were asked to distribute questionnaires to learners for completion at home; the fieldworkers in such cases simply collected completed questionnaires afterwards. This occurred in five instances.

Three schools from the original random sample of 293 secondary schools were replaced by three additional schools not included in the sample – two in Cape Town (Western Cape) and one in Pretoria (Gauteng). All three substitute schools had similar profiles in terms of geographical location and socio-economic background to the schools they replaced. Substitute schools were included primarily for three reasons:

- It was felt that the sample size should be kept at 293 schools to justify an acceptable response rate;
- The schools that refused to participate in the survey were all former white schools that were strong feeder institutions for higher education. It was deemed important to retain this profile in the sample, as there were few such units in the original sample; and
- The maintenance of a sample of 293 schools would enable the project to maximise fieldworker satisfaction levels – remuneration being linked to successful completion of fieldwork at designated schools.

4.1.3.4 Lessons learned from the fieldwork

Two of the key lessons to be drawn for subsequent fieldwork components of large-scale projects are the following:

- Adequate time (at least a six-week period) is needed to obtain provincial departmental approval for a study of this magnitude and for concomitant access to schools; and
- Notwithstanding the new democratic order in South Africa, attempts should be made, in a national study of this kind, as far as possible to match fieldworkers and questionnaire respondents in terms of racial and linguistic background. This recommendation is underpinned by evidence that:
    - A (white) project team member for whom Afrikaans is a first language enjoyed major success in persuading at least five Afrikaans-medium formerly white schools to participate in the study;
    - Many formerly white schools contacted the HSRC offices claiming not to have received any documentation pertaining to the study and / or not to have received any communication from the fieldworkers allocated to them (all fieldworkers were black) – despite HSRC evidence to the contrary; and
    - Many fieldworkers themselves (all but four of the thirty fieldworkers were African) managed to persuade African schools to participate in the study.

This last point accounts in large measure for the high level of co-operation of schools and the concomitantly high learner response rate.

4.1.4 Data analysis

Analysis of the Grade 12 Learner Choice survey data was conducted according to a highly structured analysis procedure – which is standard practice for analysing large-scale, complex survey databases. The statistical analysis procedure can be divided into two parts – descriptive analysis and inferential analysis.

4.1.4.1 Descriptive analysis

Descriptive analysis served as the most basic analytical procedure, aiming, firstly, to summarise or describe characteristics of all the variables in the Grade 12 Learner Choice survey database. From the descriptive analysis, basic statistics were obtained, such as frequencies, percentages, averages, standard deviations, etc., which provided a first-order statistical profile of learners. As this was performed for all variables, the descriptive analysis constitutes the bulk of reported data and information in this report.

The type of descriptive statistics calculated depended on the level of measurement of the variable – whether nominal, ordinal, or scale. Most variables in the Grade 12 Learner Choice questionnaire were nominal and ordinal variables. For nominal variables, frequencies and percentages were calculated, while for ordinal variables, frequencies, percentages and averages were calculated.

The second aim of the descriptive analysis was to test for possible relationships between variables by means of further descriptive analytical procedures. The decision as to which relationships to test for, out of a potentially large number of different variable permutations, was focussed on three key research questions, namely:

- Whether the learner would go to university or technikon, or not;
- Which university or technikon he / she would go to; and
- What he / she would study at university or technikon.

As most variables in the database were nominal or ordinal – that is, categorical variables – two statistical techniques suitable for categorical variable analysis were used to test for relationships. These were CHAID analysis and cross tabulation. The purpose and method of these techniques are briefly described below.

The acronym 'CHAID' stands for Automatic Interaction Detection method that uses the CHI-squared statistic. The purpose of this method is to explore the relationship between a large number of categorical variables and the outcome of a single categorical variable. The technique functions by finding and ranking all variables that have a statistically significant relationship with a single variable and can therefore be assumed to have predicted the specific outcome of the single variable. For example, an outcome of the variable 'whether to study or not' may be shown to be predicted by a set of variables such as 'income', 'gender', etc., out of a number of other variables. Results of a CHAID analysis are depicted by means of a hierarchical tree structure, of which the first branch represents the strongest predicting variable, while sub-branches represent the strongest predictors within categories of the strongest predicting variable. The CHAID technique, which is relatively new in social sciences research, has proved to be invaluable for analysing large data sets.

The second technique, cross tabulation, was used most often to determine relationships between key variables. Cross tabulations intersect two variables by listing categories of one variable in rows and categories of the other variable in columns. In such a table the particular statistic for each category of a particular variable is shown in relation to each category of another variable.

When testing for relationships between variables in the course of descriptive analytical procedures, it is important to note that results are based on a sample of learners. The key question, in the present study, was to establish whether observed relationships were true for the whole population of Grade 12 learners. For this, a second level of analysis was used, namely inferential analysis.

4.1.4.2 Inferential analysis

As indicated above, the objective in deploying inferential analysis was to establish whether observed relationships between variables from the descriptive analysis, which reflected a sample of learners, also reflected the population of Grade 12 learners in South Africa – that is, whether the relationships were *statistically significant*. Statistical inference is a statistical procedure which infers statistical results, based on a sample, to the population from which the sample is drawn. Three particular inferential statistics, called *measures of association*, were used to test relationships, namely:

- The chi-squared statistic;
- The phi-coefficient; and
- The contingency coefficient.

When results from these statistics are reported as statistically significant, it means that it is highly unlikely that the results obtained are due to some form of sampling error – that the relationship between variables, therefore, is probably an accurate reflection of the population.

CHAID analyses automatically test for significance by using chi-square statistics. All relationships highlighted by the CHAID analyses in this report can therefore be regarded as statistically significant[5] and representative of the population of Grade 12 learners. Similarly, all other relationships between variables highlighted in this report should be regarded as statistically significant.

4.1.4.3 Statistical weighting

Statistical weighting is an intentional manipulation of data in order to align the profile of the sample with the expected profile of the population. All statistics in the descriptive analysis procedure were calculated using a weighting variable. This means that the data for each learner in the database were proportionally weighted, or inflated, by a numerical factor depending on the particular sample cohort in which the learner fell. The purpose of this was to have statistics reflect on the actual population: whereas the database consists of a sample of learners, the ideal was to examine the characteristics of the entire *population* of Grade 12 learners. All statistics reported this report are, then, unless otherwise indicated, *weighted*.

The weighting variable was obtained by calculating a weight for each learner in the sample database. The weight was calculated by multiplying two weighting factors –

5 Significance in this study is calculated at the 95% confidence interval.

number of schools with number of learners per school – since the multi-stage sampling method involved, firstly, drawing a random sample of clusters (schools) from all strata (strata consisting of three categories of pass rates by nine provinces), and secondly, drawing a random sample of units (learners) from clusters (schools).

For the first sampling procedure (which involved the sampling of schools), the weight factor (W1) was calculated as follows:

$$W_1 \left(\frac{\varphi \text{ Grade 12 learners in stratum}}{\varphi \text{ Sampled schools in stratum}} \right) \text{ x } \varphi \text{ Grade 12 learners in school}$$

For the second sampling procedure (which involved the sampling of learners within schools), the weight factor (W2) was calculated as follows:

$$W_2 \left(\frac{\varphi \text{ Grade 12 learners in school}}{\varphi \text{ Sampled schools in school}} \right)$$