CHAPTER 5

RESEARCH METHODOLOGY

5.1 INTRODUCTION

Marketing research is the collecting, analysing and interpreting of marketing information pertaining to customers, competitors and the market, and thus provides information to improve marketing decision-making (Gofton & Ness, 1997:9). Consumer research, one of the focus areas of marketing research, is the gathering, recording and analysis of facts about problems relating to final households, thus the end consumer. Researching the consumer helps organisations to understand the environment, identify problems and opportunities, and develop and evaluate a marketing strategy to reach existing and potential consumers. According to McDaniel and Gates (2001:8), research relating to consumer decision-making focuses on the analysis of purchase motives, needs, buying habits, attitudes towards brands, perceptions of organisations as well as cultural and social influences. For the purpose of this study, the focus is on the consumer decision-making processes of students, specifically choice factor importance and information source usefulness, when selecting a higher education institution.

The aim of this chapter is to describe the research methodology used in this study. According to Cooper and Schindler (2003:6), the study of research methodology provides people with the knowledge and skills needed to solve problems and meet the challenges of fast-paced decision-making environments. Figure 5.1 presents the research process followed in this study and will serve as a visual reference and guide for Chapter 5.

Step 1 of the research process was addressed in Chapter 1, when the research problem or opportunity was identified and the research objectives determined. Conducting secondary research provided background information for defining the problem and setting objectives, as well as developing the questionnaire. This was addressed in Step 2 of the research process with the literature review in Chapter 2, Chapter 3 and Chapter 4.

Figure 5.1: The research process

Step 1	Identify the problem and determine the research objectives (Chapter 1 & Chapter 5, Section 5.2)
Step 2	Literature review (Chapters 2, 3 & 4)
Step 3	Set hypotheses (Chapter 1 & Chapter 5, Section 5.3)
Step 4	The research design (Chapter 5, Section 5.4)
Step 5	Develop a sampling plan (Chapter 5, Section 5.5) – Sample population (Chapter 5, Section 5.5.1) – Sample frame (Chapter 5, Section 5.5.2) – Sampling method (Chapter 5, Section 5.5.3) – Sample size (Chapter 5, Section 5.5.4)
Step 6	Select a primary research method (Chapter 5, Section 5.6) – Quantitative method: survey (Chapter 5, Section 5.6.1)
Step 7	Design the data collection instrument (Chapter 5, Section 5.7) – Developing the questionnaire (Chapter 5, Section 5.7.1) – Questionnaire lay-out and question types (Chapter 5, Section 5.7.2) – Sections of the questionnaire (Chapter 5, Section 5.7.3) – Validity and reliability (Chapter 5, Section 5.7.4) – Pre-testing (Chapter 5, Section 5.7.5)
Step 8	Gathering data (Chapter 5, Section 5.8)
Step 9	Processing data (Chapter 5, Section 5.9) – Data preparation (Chapter 5, Section 5.9.1) Editing (Chapter 5, Section 5.9.1.1) Coding (Chapter 5, Section 5.9.1.2) Data capturing (Chapter 5, Section 5.9.1.3) Data cleaning (Chapter 5, Section 5.9.1.4) – Data analysis (Chapter 5, Section 5.9.2)
Step 10	Report the research findings (Chapters 6 & 7)

Source: Adapted from Cant, Gerber-Nel, Nel, and Kotze (2003:36); Gofton and Ness (1997:14); and Welman, Kruger and Mitchell (2005).

Steps 3 to 9 of the research process will form the main focus of Chapter 5, with Step 10 of the research process being addressed in Chapter 6 and Chapter 7. In summary, Chapter 5 addresses the research process used to conduct this study by explaining the research problem, objectives, hypotheses, research design, sampling plan, data

collection methods, questionnaire design, as well as the gathering and processing of the data.

5.2 IDENTIFY THE PROBLEM AND DETERMINE THE RESEARCH OBJECTIVE (STEP 1)

South Africa's educational landscape as well as the global arena has become increasingly turbulent. Competition is much stronger than before, with more institutions joining the market (Espinoza *et al.,* 2002:19). The changing higher education landscape, nationally and globally, has encouraged the development of a market culture among higher education institutions. Institutions are now in a position where they have to compete for scare resources such as finances. To survive in this competitive environment, institutions must have an advantage, which means that a higher education institution must provide its target market with more value than its competitors. In order to provide superior value to students, higher education institutions need to anticipate and react to students' needs, thus, understanding students' behaviour. Considering all the challenges that higher education institutions are faced with, it is evident that institutions will have to become more marketing-oriented. Higher education institutions need to focus increasingly on marketing techniques used by profit organisations. One of the key issues in the successful development of a marketing strategy is to determine which factors students consider when they have to make a decision on which institution to enrol at. A proper assessment of the choice factors students consider to select a higher education institution, as well as the sources of information they make use of, will enable institutions to allocate funds, time and resources more efficiently and effectively.

Although many studies have been done internationally on choice factors considered by students in selecting a higher education institution (Espinoza *et al*., 2002; Arpan *et al*., 2003; Hoyt & Brown, 2003 and Gray & Daugherty, 2004), the authors disagree about the number of choice factors to investigate and conflicting results have been found as to the importance of the choice factors. Only a limited number of local studies addressed aspects of higher education marketing: the image of universities (De Wet, 1983 and Kruger, 1994), marketing strategies (Diederichs, 1987), market positioning

(Van Biljon, 1992), corporate image (Roux, 1994), marketing communication strategies (Jones, 2002) and corporate reputation (Coetzee & Liebenberg, 2004). However, despite these works, little is known about the choice factors and information sources considered by local students when they select a higher education institution. Furthermore, the above-mentioned local and international studies have several limitations, such as small sample sizes, investigation of a limited number of choice factors and/or information sources, and focusing on only one or two higher education institutions. Therefore, the research problem is to determine the relative importance of various choice factors and the usefulness of various information sources in the higher education institution selection process of a large sample of higher education institutions.

After identifying the research problem or opportunity, one needs to develop research objectives. A marketing research objective can be defined as the specific information needed to solve a marketing research problem (McDaniel & Gates, 2001:23). Kent (1993:320) states that research objectives are there to spell out what the research is designed to explore, measure or explain. Objectives may be spelled out in terms of a hypothesis or being formulated as questions or statements.

Literature emphasises the need for higher education institutions to identify the choice factors and various information sources used by students, in order to understand their customers better. The main goal of this study is therefore to investigate the relative importance of the choice factors, as well as the usefulness of different sources of information, considered by first year Economic and Management Sciences at higher education institutions in South Africa. By focusing on selected South African universities, the study's primary research objectives are to:

i) Determine the relative importance of each of the 23 identified choice factors that first year Economic and Management Sciences students use to select a higher education institution;
ii) Investigate the usefulness of the sources of information considered by first year Economic and Management Sciences students as perceived by ethnic groups, gender groups and academic institutions attended;

iii) Determine whether students from different ethnic backgrounds differ regarding the importance they attach to choice factors when selecting a higher education institution;

iv) Determine whether students with different home languages differ regarding the importance they attach to choice factors when selecting a higher education institution;

v) Determine whether there are gender differences regarding the importance students attach to choice factors when selecting a higher education institution;

vi) Determine whether students from different higher education institutions differ regarding the importance they attach to choice factors when selecting a higher education institution;

vii) Determine if students that live 70 kilometres or further from a higher education institution make more use of campus visits or open days as a source of information than students living close-by;

viii) Determine if students that are resident in the province in which the higher education institution is located make more use of word-of-mouth as a source of information; and

ix) Determine if students with an average of seventy percent or more in Grade 12 make more use of higher education institutions' websites than students with a lower average in Grade 12.

Step 2, the literature review, were discussed in detail in Chapters 2, 3 and 4 and will therefore not be explained again. The hypotheses that resulted from the literature review and research objectives will be explained in the next section.

5.3 SET HYPOTHESES (STEP 3)

In hypothesis testing, the aim is to examine whether a particular proposition concerning the population is likely to hold or not. The term null hypothesis reflects the concept that this is a hypothesis of no difference and therefore always includes a statement of equality. On the other hand, the alternative hypothesis is the complement of the null hypothesis and postulates some difference or inequality (Diamantopoulos & Schlegelmilch, 2000:132). An alternative hypothesis, which in addition to the existence

of a difference can also indicate the direction of the expected difference, is known as a directional hypothesis or a one-tailed hypothesis. A hypothesis that only postulates a difference without any *a priori* expectations as to the direction of the difference is called an exploratory hypothesis or a two-tailed hypothesis (Diamantopoulos & Schlegelmilch, 2000:133). Both one-tailed and two-tailed hypotheses are used in this study. The alternative hypotheses that resulted from the literature review and research objectives are explained below.

5.3.1 HYPOTHESIS 1

Bers and Galowich (2002:70) note that studies by Bodfish in 2000 and Cabera and La Nasa in 2001 suggest that the institution selection process may differ among ethnic groups. Teranishi *et al.* (2004:527) also found that the institution selection processes varied with regard to the ethnic background of students. Their study determined that factors such as cost, financial aid, perceptions of prestige and reputation of institutions differ according to ethnic group. A South African study by Cosser and Du Toit (2002:2-12), found that ethnic groups differ in the importance they attach to choice factors when it comes to higher education institution selection. Their study identified that Black African students are more influenced by parental and peer persuasion than other groups and that Black African students are more influenced by sport facilities and lower fees than the other three ethnic groups. From the above-mentioned findings in previous studies, the following hypothesis was formulated:

$H_{0:}$ Students from different ethnic backgrounds do not differ regarding the importance they attach to choice factors.

$H_{1:}$ Students from different ethnic backgrounds differ regarding the importance they attach to choice factors.

Hypothesis 1 concerns the difference between two or more groups and can be classified as a non-directional hypothesis, which is also known as a two-tailed or an exploratory hypothesis. The construct of ethnic background will be measured using a multiple choice single response question (Question 5 in the questionnaire), while the choice factors will be measured with a 5-point Likert scale containing 23 items

(Question 1 in the questionnaire). Refer to Appendix A for an example of the questionnaire and Section 5.7 for a detailed discussion of the questionnaire.

5.3.2 HYPOTHESIS 2

Language is related to ethnic groups and as mentioned above, literature suggests differences between ethnic groups and the importance of choice factors. The assumption can thus be made that possible differences can be expected between different language groups. Holtzhausen (2005:186) also found that there were differences between the importance of choice factors and language groups, indicating that high academic standards were most likely to influence English speaking respondents and that image was more important to African language respondents. It was therefore hypothesised that:

$H_{0:}$ Students speaking different home languages do not differ regarding the importance they attach to choice factors.

$H_{2:}$ Students speaking different home languages differ regarding the importance they attach to choice factors.

Hypothesis 2 concerns the difference between two or more groups and can be classified as a non-directional hypothesis. The construct of home language will be measured using a multiple choice single response question (Question 7 in the questionnaire), while choice factors will be measured with a 23-item, 5-point Likert scale (Question 1 in the questionnaire).

5.3.3 HYPOTHESIS 3

Litten (1982:390-392) found significant differences between gender and the importance these groups attach to some of the choice factors. Findings from a South African study show that female students view security as a more important choice factor than their male counterparts (Du Plooy & De Jager, 2004:227). From the findings of previous research studies, the following hypothesis was formulated:

$H_{0:}$ Male and female students do not differ in the importance they attach to choice factors.

$H_{3:}$ Male and female students differ regarding the importance they attach to choice factors.

Hypothesis 3 is a hypothesis concerning the difference between two groups and can be classified as a non-directional hypothesis. The construct of gender will be measured using a dichotomous multiple choice single response question with the options “male” or female” (Question 3 in the questionnaire), while choice factors will be measured with a 23-item, 5-point Likert scale (Question 1 in the questionnaire).

5.3.4 HYPOTHESIS 4

As mentioned in Chapter 4, Davis (1998) found that students attach high importance to factors such as the beauty of the campus, good sporting facilities and the reputation of a prestigious institution. Findings from the study by Martin *et al*. (1994:36) identified reputation as important, but added that first year students at the University of South Australia also ranked career preparation, specific academic programmes, distance from home, academic reputation, quality of research programmes and library resources as having a strong influence on their choice of university.

Interesting to note is that Bers and Galowich (2002:80) detected that factors related to money were more influential than the institution’s reputation. Geraghty (1997:41) also discovered that first year students in America place a very high importance on financial assistance and low tuition rates, as the majority of students indicated that they selected an institution for financial reasons. Findings from a study by Sevier (1993:48-50) confirmed what others found, namely that the reputation of the institution, cost and the availability of financial aid were the most important factors influencing higher education institution selection. But his research also added that the availability of study courses is important.

Coetzee and Liebenberg (2004:71) observed in a South Africa study that academic reputation, image, sporting facilities, friends studying at the higher education institution

and location were the most important choice factors. Thus, research conducted by the above-mentioned authors revealed conflicting results on the importance students attach to choice factors. The research was however conducted at different higher education institutions and it may be that the nature of the higher education institution contributed to the different importance levels reported by the respondents. It was therefore hypothesised that:

$H_{0:}$ Students from different academic institutions do not differ regarding the importance they attach to choice factors.

$H_{4:}$ Students from different academic institutions differ regarding the importance they attach to choice factors.

Hypothesis 4 concerns the difference between more than two groups and can be classified as a non-directional hypothesis. The construct of academic institution will be measured using a multiple choice single response question with the options: University of Pretoria, Tshwane University of Technology-Witbank campus, University of Johannesburg, University of the Free State, University of KwaZulu-Natal and the University of North-West (Question 6 in the questionnaire). Choice factors will be measured using a 23-item, 5-point Likert scale (Question 1 in the questionnaire).

5.3.5 HYPOTHESIS 5

Previous research findings suggest that students do not attach the same importance to the different sources of information and that there may be differences between gender, age, ethnic background, academic ability, geographical location of students and the sources of information that they use (Martin, 1994:28-36; Kotler & Fox, 1995:259; Hamrick & Stage, 2004:151).

Open days and career exhibitions were identified as the most important sources of information used by prospective students (Coetzee & Liebenberg, 2004:71). The authors also determined that prospective students attach high importance to activities on campus as a source of information. Students furthermore indicated that they are

more willing to attend an exhibition or open day to secure firsthand information than to rely on mass media.

Martin (1994:37) detected that geographical location of students also influenced their institution selection process. Brochures and websites do not provide enough information and students living far from a university have to do their own research by visiting the campus. It was found that these students usually make an effort to attend a campus visit to gain firsthand experience (Anon, 2006c and Evans, 2006). It could thus be assumed that a campus visit or open day is the perfect way to display the campus to prospective students living far away (Anon, 2004b). From the afore-mentioned it can be hypothesised:

$H_{0:}$ Students that live further than seventy kilometres from the higher education institution do not attach more value to campus visits or open days as a source of information than students living close-by.

$H_{5:}$ Students that live further than seventy kilometres from a higher education institution attach more value to campus visits or open days as a source of information than students living close-by.

Hypothesis 7 concerns the difference between two groups and can be classified as a directional or one-tailed hypothesis. The sources of information (open days and campus visits) will be measured with a 5-point Likert scale (Question 2.4 in the questionnaire). The distance from parents' homes to university attended will be measured using a multiple choice single response question (Question 8 in the questionnaire).

5.3.6 HYPOTHESIS 6

Personal sources of information refer to information shared by friends, family or reference groups and is known as word-of-mouth. Zeithaml and Bitner (2000:32-33) state that the service industry recognises the strong influence of word-of-mouth. Research shows that when purchasing a service product such as education, students rely to a greater extent on personal sources, such as word-of-mouth by friends, family,

opinion leaders and teachers, because mass media convey very little about experience qualities (Jones, 2002:76). Findings by Cosser and Du Toit (2002:2) suggest that the family has a high degree of influence in the institution selection process, especially by encouraging students to continue to study. Word-of-mouth information, in particular, is regarded as highly credible and believable. Cosser and Du Toit (2002:101-103) also observe that discussion with relatives and friends are well used by high school learners as sources of information. It is therefore hypothesised that:

$H_{0:}$ Students who are residents of the province in which the chosen higher education institution is located, do not value word-of-mouth (friends and other people) more as a source of information than students from other provinces.

$H_{6:}$ Students who are residents of the province in which the chosen higher education institution is located, value word-of-mouth (friends and other people) more as a source of information than students from other provinces.

Hypothesis 6 concerns the difference between two groups and can be classified as a directional or one-tailed hypothesis. The sources of information (word-of-mouth) will be measured with a 5-point Likert scale (Question 2.8 in the questionnaire). Residence in the province in which the university is located will be measured with a dichotomous multiple choice single response question with options “yes” or “no” (Question 9 in the questionnaire).

5.3.7 HYPOTHESIS 7

Websites were identified as one of the most important sources of information used by South African students (Coetzee & Liebenberg, 2004:70). Cosser and Du Toit (2002:101-103) found that receiving information directly from a university was the most important source of information used by high school learners. Before a higher education institution can decide what content is essential for its website, it should determine the target audience, their needs and demographic profile (Joch, 2006). The latest South African web user survey (Anon, 2005b) indicates that the typical South African web user have either completed university or have a form of post matric qualification, indicating a high scholastical aptitude. Karuppan (2001:138-149)

determined that scholastic aptitude or academic ability of students influenced their web usage, as students with a higher academic ability made more use of the web. From the afore-mentioned it can be hypothesised that:

$H_{0:}$ Students with a Grade 12 average of seventy percent or more do not rely significantly more on higher education institution's websites as a source of information than students with a lower Grade 12 average.

$H_{7:}$ Students with a Grade 12 average of seventy percent or more rely significantly more on higher education institution's websites as a source of information than students with a lower Grade 12 average.

Hypothesis 7 concerns the difference between two groups and can be classified as a directional hypothesis. The sources of information (websites) will be measured with a 5-point Likert scale (Question 2.3 in the questionnaire). The average grade for students in his/her final Grade 12 exam will be measured using a multiple choice single response question (Question 10 in the questionnaire). Answers to this question will be collapsed to form 2 groups: "Grade 12 average of more than 70 percent" (higher academic ability) and a second group with a "Grade 12 average of lower than 70 percent" (lower academic ability). This was done to ensure equal cell sizes as well as deemed to be an appropriate split for academic performance.

Steps 1 and 3 of the research process addressed the research problem, the research objectives and the hypotheses. These steps determined to a large extent the kind of research design that is most suitable for the study and forms the basis for the next section. Section 5.4 will describe the research design followed for this study as part of Step 4 in the research process.

5.4 THE RESEARCH DESIGN (STEP 4)

The research design delineates the structure of the investigation in such a way as to attain answers to the research objectives. A research design is thus a preliminary plan for conducting research. According to Cooper and Schindler (2003:146-151), a broad research design consists of different elements, such as the type of research design,

focus, time dimension and conditions. These elements, as applied to the study at hand, are briefly discussed below.

In this study, a formal research design was used to test the stated hypotheses. This quantitative study followed a two-stage research design. The first stage included an extensive literature review of secondary data on the higher education landscape globally and locally, the marketing of higher education institutions and the consumer decision-making process students go through when selecting a higher education institution. The literature review consisted of journal articles, research reports, Internet searches and textbooks, covering the following range of disciplines: marketing, consumer behaviour, the higher education landscape and research methodology. According to McDaniel and Gates (2001:76), it is always advisable to do desk research, which entails the scanning of all available secondary data sources before engaging in primary data search. This approach guides the researcher in identifying unanswered questions and prevents him/her from replicating research without the necessary background knowledge on what has already been achieved.

The second stage was a formal descriptive study that was conducted to investigate the objectives previously mentioned. In this stage, students were questioned and their responses collected by means of a structured questionnaire. The researcher executed an ex post facto design; this implies that the researcher had no control over the variables in the sense of manipulating them. The researcher was only able to describe/report the importance of the choice factors in the institution selection process and other related issues.

The research design has a descriptive focus, as the purpose was to determine the relative importance of each choice factor and the usefulness of various sources of information. The purpose of descriptive research is to describe a problem or opportunity by providing a broad overview of a sample of a large population (Mouton, 2001:152). According to Churchill and Peter (1998:65), this type of research is helpful to describe the characteristics of certain groups, to estimate the proportion of people who behave in a certain way and to make specific predictions. This study aimed to provide a description (in terms of convenience sample) of the following:

- Socio-demographic information: the study provides a brief profile of first year Economic and Management Sciences students in terms of gender, age, ethnic background, higher education institutions attended, home language, proximity of homes from the higher education institutions and academic standing.
- Behavioural information: the decision-making process (specifically the choice factors and information sources influencing institution selection) of first year students.

The time dimension of the study was cross-sectional, as the data collected represents a snapshot of one point in time. Due to the fact that this study focuses on the breadth rather than depth of the study, it could be classified as a statistical study. Research was conducted under field conditions, as students were interviewed in their lecture rooms at the higher education institution they were attending. Participants (students) were fully aware of the research activity, which should not influence the results, as they have nothing to lose or gain from manipulating their answers.

After designing a research plan it is necessary to decide which respondents to include in the research study. Step 5 of the research process focuses on the development of a sampling plan and is addressed in the next section.

5.5 DEVELOPMENT OF A SAMPLING PLAN (STEP 5)

Sampling, which involves selecting a subset of the population, was chosen above a census of the whole population, due to time and cost constraints. According to Schmidt and Hollensen (2006:159-160), sampling involves the identification of a group of individuals or households who can be reached by mail, by telephone or in person and who possess information relevant to solving the marketing problem at hand. Lamb *et al*. (2004:265) define a sample as a subset of the target population from which information is gathered to estimate something about the population. The process of developing a sampling plan consists of 5 steps: identifying the sample population, identifying the sample frame, selecting a sampling method, selecting the sample size and gathering information from the sample elements (Lancaster, Withey & Ashford, 2001:40). These steps will now be briefly described.

5.5.1 SAMPLE POPULATION

A population can be described as the entire compilation of elements that the researcher aspire to draw conclusions from (Cooper & Schindler, 2003:179). Webb (2002:48) points out that a population or universe is the aggregate of all the elements.

As Grade 11 and Grade 12 learners could not be used as a sample population due to the difficulty in obtaining permission from the Department of Education, the target population for this study was first year Economic and Management Sciences students at six higher education institutions in five provinces, namely the University of Pretoria (Gauteng), Tshwane University of Technology-Witbank campus (Mpumalanga), University of Johannesburg (Gauteng), University of the Free State (Free State), University of KwaZulu-Natal (KwaZulu-Natal) and North–West University (North-West). These six universities represent some of the major universities in South Africa situated in five different provinces. It should be noted that the Tshwane University of Technology-Witbank campus was specifically included in sample and not the main campus in Tshwane, as there are no other major higher education institutions in the Mpumalanga pronvince. Unfortunately, the major universities in the Western and Eastern Cape did not want to partake in the study. The research was conducted in February, March and April of 2006. These months were chosen as it is in the beginning of the year and to ensure that the first year students were still relatively uncontaminated and could still remember how they made their selection of a higher education institution the previous year.

After determining the sample population of the research study, a list or sample frame containing the identified members of the population must be obtained.

5.5.2 SAMPLE FRAME

A sample frame is a list of population members used to obtain a sample (Aaker, Kumar, Day & Lawley, 2000:367). A frame may be a register of industries, a telephone directory or even a map. Hair, Bush and Ortinau (1996:330) state that a sample frame must meet the following requirements:

It must represent all the elements of the population;

- There must be no duplication of elements; and
- It must be free from foreign elements.

It was not possible to obtain up-to-date lists from all six higher education institutions participating in this study to be used as sample frames. Thus, for the purpose of this study, the sample frame was broadly defined as first year Economic and Management Sciences students attending lectures.

The next step involves deciding how to select some elements of the sample population by making use of the sample frame, thus, choosing a sampling method.

5.5.3 SAMPLING METHOD

There are two approaches to sampling: probability and non-probability sampling. Probability sampling is based on the premise that each member of the population has a definite opportunity to be selected. With non-probability sampling, there is no guarantee that each member of the population has a definite opportunity to be selected (Diamantopoulos & Schlegelmich, 2000:11-13). According to Cooper and Schindler (2003:200) and Shao (1999:38), non-probability sampling is more cost-effective, faster and effective.

Non-probability sampling was chosen, as the characteristics of this method have particular appeal due to financial and time restraints. According to McDaniel and Gates (2001:336), non-probability samples can produce samples of the population that are reasonably representative. A major implication of this type of sampling is that a statistical evaluation of the sampling error cannot be undertaken (Tustin *et al.,* 2005:344). Measurement errors, which are usually associated with non-probability sampling, were prevented in the following ways: questionnaire design errors were prevented by following the guidelines for constructing a questionnaire and by using focus groups as a pilot study to test the questionnaire; cross-checking, computer checks and pre-coding were used to avoid coding and data capturing errors; respondents' errors were addressed by making sure it is not a lengthy questionnaire; and ego/humility questions were minimised. Interviewer errors were limited by making

use of a structured questionnaire. Non-response errors can also affect the results and therefore students completed the questionnaire during class time while their lecturer was present, to increase the response rate and to ensure the correct elements were included in the sample.

Three different types of non-probability sampling methods can be distinguished: judgmental, quota and convenience sampling (Welman, *et al.,* 2005:56). For the purpose of this study, a convenience sample was chosen. Hair *et al.* (2000:354) define convenient sampling as a study in which respondent participation is voluntary or which leaves the selection of sampling units primarily up to the interviewer. Thus, convenience sampling was used in this study as the sample members (first year Economic and Management Sciences students) were chosen on the basis of being available or accessible during normal class times. To avoid the potential bias owing to the use of non-probability sampling, the questionnaire was administrated at six different higher education institutions, at different times and classes by different lecturers who acted as fieldworkers/interviewers. Due to the fact that this sample was drawn using a non-probability method, the researcher will not be able to assess the likeliness that the sample is unrepresentative and by how much (Diamantopoulos & Schlegelmilch, 2000:13).

After selecting a sampling method, the sample size was determined.

5.5.4 SAMPLE SIZE

The sample size is determined by both statistical and practical considerations (Jarboe, 1999:38). Tustin *et al.* (2005:361) state that statistical calculations of sample sizes can only be done for probability sampling methods, as there is no statistical formula for prior calculation of the size of a non-probability sample. As this study made use of a non-probability sampling method, the sample size was determined by practical concerns such as resources in terms of money, time and personnel impact, as well as the number of respondents present in class and willing to participate. Keeping the latter in mind and to prevent the non-response error associated with non-probability sampling, as well as to ensure that group comparisons could be made when analysing

the data, 250 questionnaires were distributed to each of the six higher education institutions, thus a total of 1500 questionnaires.

The next section will explain the primary data collection method used to collect the data from the respondents.

5.6 SELECT A PRIMARY DATA COLLECTION METHOD (STEP 6)

Data collection refers to the way in which the data was captured in the field setting. According to Martins *et al*. (1996:122), primary data is obtained through qualitative and quantitative methods. Qualitative research methods intend to gather in-depth, detailed information through methods such as in-depth interviews, projective techniques and focus groups (Welman *et al.,* 2005:188). Quantitative research methods focus on gathering a large amount of information through surveys such as mail, telephone and personal interviews. Qualitative methods are more unstructured, flexible and diagnostic than quantitative methods, and aim to obtain information from respondents in an indirect manner. These methods are useful in exploratory research and are appropriate for hypotheses generation (Du Plessis & Rousseau, 2005:28).

Quantitative methods are systematic and structured, and aimed at obtaining information from respondents in a direct, open manner. Results obtained from these methods are easily quantifiable and have a potentially high degree of accuracy. These methods are often used for testing hypotheses (Du Plessis & Rousseau, 2005:22). The primary advantage of quantitative methods is that a great amount of data can be collected about the individual respondent at one point in time. These methods are also very versatile, because they can be applied in virtually any setting and can be adapted for almost any research objective (Barrie & Furnham, 1992:20; Martins, Loubser & Van Wyk, 1996:140 and Jarboe, 1999:51).

Section 5.6.1 will explain the quantitative method used in the study.

5.6.1 QUANTITATIVE METHOD: SURVEY

There are many different survey methods, such as personal interviews, telephone interviews and mail surveys. Personal interviews involve some level of face-to-face contact between the interviewer and the respondents. Based on the means of contacting respondents as well as the level of interviewer involvement, several types of personal interview methods can be distinguished: door-to-door interviews, executive interviews, shopping centre intercepts, purchase intercept techniques, omnibus surveys and self-administrated interviews. For the purpose of this study, the self-administrated interview method was used. In the self-administrated interview method, no interviewer is involved in completing the self-administrated questionnaire (Aaker, Kumar & Day, 2005:179 -181).

Tustin *et al*. (2005:184) define a self-administrated questionnaire (refer to Section 5.7) as a traditional paper questionnaire used for surveys. Self-administrated surveys allow respondents to complete the questionnaires themselves. Data collection through such written communication requires respondents to record their response to the research questions in writing. Although this method reduces cost as no interviewers are needed, there are some disadvantages, such as a high non-response rate, respondents not understanding some questions, and no interviewers present to clarify. Preventative measures were followed to minimise the problems associated with self-administrated questionnaires by ensuring that lecturers acted as interviewers and that they were in class to assist, provide clarification and motivate students (respondents) to complete the questionnaires. Punch (2003:42) suggests that researchers must strive for a response rate of at least 60 percent. The response rate of this study was 83 percent, as 1241 of the 1500 questionnaires were completed and returned. There was control over who completed the questionnaires, as the designated lecturers at each institution acted as fieldworkers. Because no mailing was involved, no up-to-date mailing list was needed and results were obtained relatively fast, as questionnaires were handed in directly after completion in class. The survey was less impersonal than a mail survey, as the lecturers were present to explain the research procedure and to provide assistance in case of any problems. This method of data collection was used as it is relatively fast, cheap, convenient, and easy to process and the response rate is usually high.

The data collection instrument, namely a questionnaire, which was used in the self-administrated interview, will be explained in detail in the section below in Step 7 of the research process.

5.7 DESIGN OF THE DATA COLLECTION INSTRUMENT (STEP 7)

McDaniel and Gates (2001:289) define a questionnaire as a set of questions designed to generate the data necessary to accomplish the objectives of the research project. It is guided by the research questions and serves as a data collection tool (Punch, 2003:30). A questionnaire provides standardisation and uniformity in the data gathering process. It standardises the wording and sequence of questions and ensures that every respondent sees and hears the same words, and every interviewer asks identical questions. It can therefore be seen as a control device. According to Webb (2002:89), a questionnaire is designed for four purposes:

- To maximise the accuracy and relevancy of information to be obtained;
- To maximise the participation of relevant elements in the sample;
- To facilitate the gathering of information; and
- To meet research objectives.

Keeping the above-mentioned in mind, the questionnaire was developed based on two existing questionnaires.

5.7.1 DEVELOPING THE QUESTIONNAIRE

According to Hoyt and Brown (2003:3), higher education institutions may develop their own in-house surveys or use a standardised instrument. The Admitted Student Questionnaire (ASQ) and Cooperative Institutional Research Programme Freshmen Survey (CIRP) were used to compile the measurement instrument for this study.

5.7.1.1 The Admitted Student Questionnaire

The Admitted Student Questionnaire (ASQ) admitted students to tell institutions what they really think of their programmes, recruitment literature, financial aid packages, competition, the institution's image and characteristics. The ASQ provides accurate information for detailed market research and data analysis (College Board, 2005). The ASQ questionnaire consists of 92 questions and is very popular internationally, as it enables institutions to find out what students think about their specific higher education institution in general as well as about specific aspects of the university experience, both academically and recreational. The advantage of using the ASQ is that it is cost-effective, convenient, accurate and effective (College Board, 2005). The added advantage is that institutions can customise the ASQ by adding their own characteristics or choice factors. The ASQ questionnaire can also be integrated with enrolment planning services, recruitment planning and financial aid programmes (College Board, 2005).

The importance of the institutional characteristics in the ASQ is measured with a 3-point Likert scale (very important, somewhat important and not important). The majority of the factors influencing student decision-making measured in the ASQ were used in Section A of the questionnaire for this study. The corresponding question of the questionnaire used for this study, is indicated in brackets. Some of the factors were however combined in one question in the instances where pre-testing revealed that students perceived them as being almost identical. The 20 institutional characteristics (choice factors) in the ASQ are listed below.

- Quality of faculty (Question 1.2);
- Quality of majors of interest to you (Question 1.2);
- Overall academic reputation (Question 1.15);
- Quality of academic facilities (Question 1.3);
- Variety of courses (Question 1.1);
- Cost (Question 1.5);
- Prominent intercollegiate athletics (Question 1.7);
- Quality of social life (Question 1.8);

- Surroundings (Question 1.10);
- Part of the country in which the college is located (Question 1.6);
- Attractiveness of campus (Question 1.9);
- Quality of on-campus housing (Question 1.11);
- Chance to be with students from different backgrounds (Question 1.19);
- Ease of getting home (Question 1.6);
- Athletic programmes in which students would like to participate (Question 1.7);
- Access to faculty (Question 1.4);
- Availability of extracurricular activities (not used);
- Access to off-campus culture and recreational opportunities (not used);
- Availability of religious activities (not used); and
- Concentration on undergraduate education (not used).

Various university characteristics/choice factors were not included and used for the purposes of this study. The choice factor “concentration on undergraduate education” was not included in the survey as the higher education institutions that participated in this study anyway focus on undergraduate education, thus making it irrelevant to include. Another factor namely “religious activities” were also not included due to its sensitive nature as well as its limited relevance for this study. The factors “availability of extracurricular activities” and “off-campus culture” were also not included due to the length constraints of the questionnaire.

Also included in the ASQ are questions on the importance of other people’s opinions in making a college choice, such as parents (Question 2.6), guidance counsellors (Question 2.7), high school teachers (Question 2.7), friends (Question 2.8), graduate and professional schools (Question 3.8), and potential for future employers (Question 1.21). All of the above-mentioned were questions used in either Section A or Section B of the questionnaire, as indicated in brackets.

The quality of information made available to students compared to other institutions is also rated in the ASQ on a 5-point Likert scale and include sources such as (the corresponding question of the questionnaire used for this study is indicated in brackets):

- Visits by admission staff to the high school (Question 3.1);
- College publications (Question 3.2);
- Communication about financial aid (Question 3.2);
- College websites (Question 3.3);
- Electronic communication (Question 3.3);
- Campus visits (Question 3.4);
- On-campus admission interview (Question 3.4);
- Contact with colleges after you were admitted (Question 3.4);
- Contact with faculty from the college (Question 3.4);
- Contact with students who attend the college (Question 3.5);
- Contact with graduates of the college (Question 3.8);
- College sponsored meetings in students' home area (not used);
- Contact with coaches (not used); and
- College's video or CD-ROMs (not used).

The majority of information sources used in the ASQ were used to develop Section B of the questionnaire used in this study. Some sources of the ASQ were combined after pre-testing indicated that South African students perceived some sources as being the same. The remaining sources such as videos and CD ROMs, college sponsored meetings in students' home area and contact with coaches are not really used by the South African higher education institutions that participated in this study and were therefore not included.

The remainder of the ASQ questionnaire deals with the image of the institution (Question 1.23), other institutions students applied for as well as college cost and financial aid (not in the scope of this study). The ASQ final section obtains information on socio-demographics such as:

- Gender (Question 3);
- Ethnic orientation (Question 5);
- Distance from college to students' homes (Question 8);
- Resident of the state in which the higher education institution attended is located (Question 9);
- Average grade in high school (Question 10);
- Type of high school attended (not included); and
- Income of parents (not included).

All the socio-demographic questions of the ASQ were included in Section C of the questionnaire used for this study, except the type of school attended and income of parents. The latter was not relevant to the study and sensitive by nature and the type of school attended did not fall within the scope of this study.

Goebel and Simonetta (2005) criticised the ASQ and advised institutions not to use the ASQ as their only recruitment and marketing research tool. They state that the biggest flaw is that it is a mail survey and therefore response rates and biases, respondent screening, questionnaire integrity, and sponsor identity problems may occur. However, they still state that it can be of good use especially for longitudinal analysis to measure how perceptions and behaviour of students may change over time. This study however used the ASQ as a self-administrated survey and not as a mail survey and therefore prevented the possible errors stipulated by Goebel and Simonetta.

5.7.1.2 The Cooperative Institutional Research Program (CIRP)

The Cooperative Institutional Research Program (CIRP) assist all stakeholders in higher education research in understanding the interest values, attitude and self-reported abilities of incoming first year students. Information from the CIRP aids institutions in decision-making, assessment activities, accreditation activities and the development and assessment of an institution's educational programme. The student profile provided by CIRP can be used by various departments of the institution to communicate, guide and modify student recruitment material and academic and support programmes (Randall, 2001). The CIRP is a national longitudinal study of the

American higher education system. Established in 1966, CIRP is America's largest and oldest empirical study of higher education, involving data on 1 800 institutions and over 11 million students. It is regarded as the most comprehensive source of information on college students (Donta, 2005). The four page questionnaire, containing 40 questions, covers a broad array of issues and themes listed below:

- Socio-demographics, such as gender, age, language, average grade in high school, year graduated, resident of the state in which the higher education institution is located, distance from college to students' home and type of high school attended, income of parents and parents' occupations. As the majority of socio-demographic questions overlapped with the ASQ, only age (Question 4) and language (Question 7) were added from the CIRP to Section C of the questionnaire. Due to the fact that six higher education institutions participated in the study, Question 6 was added to identify the different higher education institutions;
- Expectations of college (not in scope of study);
- High school experiences (not in scope of study);
- Degree goals and career plans (not in scope of study);
- College finances (not in scope of study);
- Attitudes, values, life goals (not in scope of study); and
- Reasons for attending a specific college.

The last section of the CIRP, which measures the importance of different reasons influencing the decision-making process, was of special importance, since this is relevant to the study at hand and is measured with a 3-item Likert scale (very important, somewhat important and not important). It included the following factors:

- My relatives wanted me to come here (Question 1.13);
- My parents wanted me to come (Question 1.12 / Question 2.6);
- My teacher advised me (Question 2.7);
- This college has a good reputation for its social activities (Question 1.8);
- I was offered financial assistance (Question 1.16);
- The cost of attending this college (Question 1.5);

- High school counsellor advised me (Question 2.7);
- I wanted to live near home (Question 1.6);
- Financial aid not offered by first choice (not included in the study);
- This college's graduates get good jobs (Question 1.21);
- This college's graduates gain admission into top graduate schools (not included in the study);
- I was attracted by the religious orientation of the college (not included in the study);
- I wanted to go to a school about this size (not included);
- Rankings in national magazines (Question 2.11);
- Information from a website (Question 2.3); and
- A visit to the campus (Question 2.4).

These reasons were included in Section A (Question 1) and Section B (Question 2) of the questionnaire, as indicated in brackets.

As indicated in the above-mentioned discussion of the ASQ and CIRP, the majority of choice factors and the sources of information and socio-demographics of the questionnaire for this study were based on these two questionnaires. As the ASQ especially make provision for institutions to add additional choice factors, the following factors were added, based on findings from other studies: safety and security (refer to Section 4.8), international links (refer to Section 4.8), language polices (Section 2.3.3) and flexible study mode (Section 2.3.3). Section B comprised of the information sources identified in the above discussion from the ASQ and CIRP, as well as advertisements on radio, television, magazine and campus events. These additional sources were identified by other studies (refer to Section 4.7.1) and confirmed in the pre-testing of the questionnaire as being used by students. It was also decided to make use of a 5-point Likert scale instead of a 3-point Likert scale for uniformity, as some components such as information sources in the ASQ were measured with a 5-point Likert scale. A further benefit of a 5-point Likert scale is that it provides a higher level of intensity and thus provides a finer measurement.

After the basic questions were formulated, special attention was given to the lay-out of the questionnaire as well as the different types of questions.

5.7.2 QUESTIONNAIRE LAY-OUT AND QUESTION TYPES

Lamb *et al*. (2004:262) state that by arranging the questions logically and observing other sequencing rules, the researcher enhances the standard of the interview, helps the interviewer and induces a logical flow through the questions. The following principles were followed in the design of the questionnaire for this study:

- The first question was easy and interesting to motivate the respondent to react;
- A logical sequence of questions was used;
- Sensitive questions, questions on embarrassing subjects, and demographics were positioned close to the end of the questionnaire; and
- It was ensured that a structured answer to a question does not provide respondents with the answer to other questions.

The criteria specified to ensure a good questionnaire were also applied in the study and are discussed below (Baker, 1999:162).

- It should provide the necessary decision-making information.
  Any questionnaire that fails to provide important insight for management in decision-making information should be disregarded or revised. In this study, questions were grouped under three sections: higher education institution characteristics, information sources and personal profile. This ensured that the necessary information was obtained to reach the objectives of the study and to enable higher education institutions to make the correct decisions regarding the student market.

- It should fit the respondents' requirements.
  When designing the questionnaire, the researcher must consider the topic, the type of respondent, interviewing environment and length of the interview to ensure that respondents will give their cooperation. In this study, pre-testing of

the questionnaire was used and the necessary changes implemented to ensure that the respondents' requirements were met.

- It should meet editing, coding and data processing requirements.
  The questionnaire was assigned codes before distribution and thoroughly discussed with the data capturer and statistical analyst to ensure that all the requirements were met.

According to Schmidt and Hollensen (2006:151), two types of questions can be distinguished. The first type is open response or structured questions with unstructured responses, which allow the respondent to give his/her own answer in his/her own way. Webb (2002:99) points out that this type of question forces the respondent to think, which is ideal in situations where all possible answers to a given question are not known. But they elicit much irrelevant information, lengthen the fieldwork and make coding and processing more difficult as well as confusing the students; therefore they were used to a lesser extent in the questionnaire. Open-ended questions were only used in instances where respondents wanted to indicate "other" responses that were not specified in the question and to indicate their age in years.

The second type of questions is close-ended or structured questions with structured responses. Questions of this type have various possible answers from which the respondent has to pick one or more (Martins *et al.,* 1996:221). Included in this category are dichotomous, multiple choice, closed-ended questions with single or multiple answers, as well as scaled questions. According to Jarboe (1999:70), the advantages are the same regardless of the type of closed response format used. Such questions are easier to answer, require less effort by the researcher and make tabulation and analysis easier. The answers of these questions are also directly comparable from one respondent to another and therefore mainly closed response format questions were used in the study.

The three sections of the questionnaire and the different questions of each section will be explained below.

5.7.3 SECTIONS OF THE QUESTIONNAIRE USED IN THIS STUDY

As mentioned in Section 5.7.1, the ASQ and the CIRP questionnaires were used as the main basis for the questionnaire in this study and adapted to South African students for the purpose of this study. The questionnaire consisted of three sections and each of these sections and questions in the questionnaire will be explained below (Refer to Appendix A for the complete questionnaire).

5.7.3.1 Section A of the questionnaire

Section A consisted of Question 1 and investigated university characteristics/choice factors. Question 1 contained twenty-three variables (V1-V23) and measured the importance of higher education institution characteristics (choice factors) such as:

- Wide choice of subjects/courses;
- Quality of teaching;
- Academic facilities (libraries and laboratories);
- Entry requirements;
- Fees (cost);
- Location of the higher education institutions;
- Sport programmes;
- Social life on campus (Rag, music festivals, campus dances);
- Attractiveness of campus;
- Campus safety and security;
- On-campus housing/hostels;
- Parents went there (tradition);
- Brother/sister went there;
- Friends went there;
- Academic reputation (prestige);
- Financial assistance (bursary and loans);
- Language policy;
- Links with the industry;
- Multi-culturality/diversity;

- International links (study and job opportunities);
- Employment prospects (possible job opportunities);
- Flexible study mode (evening classes and use of computers); and
- The image of higher education institutions.

A 5-point Likert scale was used to measure the 23 identified choice factors. This scale type is a widely used rating scale that requires the respondents to indicate the degree of agreement or disagreement with each of a series of statements about an object (Schmidt & Hollensen, 2006:120). Typically, each scale item has 5 response categories ranging from not important at all, of little importance, moderately important, very important, to extremely important.

Example:

	Not important at all	Of little importance	Moderately important	Very Important	Extremely Important
Factors considered	**1**	**2**	**3**	**4**	**5**
1. Wide choice of subjects/courses	1	2	3	4	5

5.7.3.2 Section B of the questionnaire

Section B of the questionnaire consisted of Question 2 (V24-V35), which measured the usefulness of different information sources such as:

- School visits by higher education institution staff;
- Higher education institution publications (newsletters and brochures);
- Higher education institution website;
- Campus visits and open days;
- Alumni;
- Parents;
- High school teachers;
- Word-of-mouth (friends and other people);

- Advertisements on radio;
- Events on campus (music festivals, Rag, sports events);
- Advertisement in magazines/newspapers; and
- Advertisements on television.

A 5-point Likert scale was used to measure the usefulness of information sources. Each scale item had 5 response categories ranging from very poor, poor, fair, good to excellent. Provision was also made for not applicable to ensure that respondents were not forced to rank an information source that they had not encountered before.

Example:

	Very Poor	Poor	Fair	Good	Excellent	Not Applicable
Usefulness of information sources	**1**	**2**	**3**	**4**	**5**	**0**
1. School visits by university staff	1	2	3	4	5	0
2. University publications (newsletters & brochures)	1	2	3	4	5	0
3. University web site	1	2	3	4	5	0

5.7.3.3 Section C of the questionnaire

Section C of the questionnaire contained Questions 3 to 10 and aimed to measure the socio-demographic details of respondents such as: age, gender, ethnic background, higher education institutions attending, home language, residence of province, proximity of home from higher education institutions and academic standing. In this section, mainly dichotomous and multiple choice questions with single answers were used. Dichotomous questions were used in Question 3 and Question 9 and provided the respondent with only two options to choose from, for example:

Are you a resident of the province in which this higher education institution is located?

Yes	***1***
No	***2***

Multiple choice questions with a single answer were used in Questions 5, 6, 7, 8 and 10.

Example:

Indicate your ethnic background

Black African	***1***
Coloured	***2***
Indian	***3***
Caucasian	***4***
Other	***5***

The questionnaire was pre-tested to uncover possible problem areas before using the instrument to collect data.

5.7.4 VALIDITY AND RELIABILITY OF QUESTIONNAIRE

A good and fair measurement tool must always adhere to the criteria of being reliable and valid. Reliability measures the accuracy and precision of the tool; thus an index that registers the extent to which measured data is free from random error (Cooper & Schindler, 2003:231). Reliability refers to the ability of a scale to produce a consistent result if repeated measurements are taken. For the purpose of determining the reliability of this study, Cronbach alpha testing will be used with a 0.7 cut-off point (Nunnally, 1987). Reliability is a necessary condition for quality measurement, but not sufficient if done alone. Before accepting and using any measure, one must also ensure its validity.

Validity refers to the accuracy of the tool actually measuring what it is supposed to measure and to the extent to which a particular measure is free from systematic and random errors (Diamantopoulos & Schlegelmilch, 2000:33). The questionnaire for the study was developed using choice factors from similar studies and two widely used international instruments (the ASQ and CIRP) as a point of reference and adapted to the South African context. The questionnaire used in this study is appropriate to the research problem and purpose and has been scrutinised by a panel of experts. The questionnaire makes use of categories that will best partition the data for testing hypotheses and showing relationships. The validity of the measurement instrument

used in this study is supported by the above-mentioned subjective arguments and judgments and is referred to as face validity or consensus validity (Aaker *et al*., 2005:208).

5.7.5 PRE-TESTING THE QUESTIONNAIRE

Pre-testing is a strong recommendation in order to confirm that the considerations and criteria stated in Section 5.7.2 have been correctly implemented and to make adjustments if necessary (Strydom *et al.,* 2000:156). The questionnaire was pre-tested by three focus groups that were conducted in December 2005. The first focus group was conducted at the University of Pretoria consisting of eight Caucasian first year students with an equal amount of male and female students. The second focus group was also conducted at the University of Pretoria and consisted of six Caucasian first year students - four male and two female. The third focus group was conducted at the Tshwane University of Technology's Witbank campus and consisted of ten Black African first year students, of which seven were female and three male. The focus groups were conducted by a moderator and the discussion was relatively structured as choice factors and information sources identified by other studies such as the ASQ and CIRP were used as a base of reference and questioning.

The results obtained in the focus group interviews were used in the refinement of the self-administrated questionnaire. Pre-testing further eliminated the number of "other" responses and ensured that the classification set was not too limited. Pre-testing the questionnaire also ensured that a specific answer could only fit into one cell of a category set; thus be mutually exclusive. Pre-testing the questionnaire also gave the researcher the opportunity to discover how the respondents reacted to the questions, ensure that the language was understandable and meaningful, that there was continuity and flow in the questions, that the respondents understood the questions and were able to complete the questionnaire, and to determine if the questionnaire was not too long and time consuming (Punch, 2003:34). The pre-test indicated that the questionnaire took the average student 10-20 minutes to complete. Students also identified unfamiliar words, abbreviations and instructions. The necessary language changes were made to the questionnaire and clear instructions added to ensure that students would be able to understand and respond to each question in the

questionnaire properly. The necessary changes and improvements were made before fieldwork commenced.

The questionnaire was used to collect the data from the sample. The data collection process will be addressed in the next section as Step 8 in the research process.

5.8 GATHERING DATA (STEP 8)

Gathering data is a two-stage process. First, the sample elements must be selected and secondly information must be gained from those elements (Schmidt & Hollensen, 2006:170). The sample elements refer to the subjects on whom the measurement was undertaken (Cooper & Schindler, 2003:179).

Before the data collection could proceed, permission had to be obtained from the higher education institutions. Six of the eight institutions approached agreed to participate in the study and a designated lecturer at each institution was identified, who acted as a contact person and fieldworker. The researcher gave a brief description of the research project, reasons for conducting the research and timeframe to each fieldworker telephonically. A total of 250 questionnaires were then sent to each institution. The fieldworker at each institution had four weeks to arrange a convenient time to distribute the questionnaires. Written permission (by means of consent forms) was also obtained from each respondent (students) to participate in the study. The front page of the questionnaire contained instructions on how to complete the questionnaire as well as clear instructions at each question. The fieldworker was available for any questions or uncertainties and to ensure that only first year Economic and Management Sciences students completed the questionnaire. After completion, the questionnaires were handed back to the fieldworker, who then sent it back to the researcher. The research was conducted in February, March and April of 2006. These months were chosen because it is in the beginning of the year, and to ensure that the first year students were still relatively uncontaminated and could still remember how they made their selection of a higher education institution the previous year.

After data is collected from respondents, data must be processed. Step 9 in the research process describes the data processing procedures and will be discussed below.

5.9 DATA PROCESSING (STEP 9)

After data has been collected, it has to be processed. This consists of data preparation and data analysis. Both will be described in the subsequent sections.

5.9.1 DATA PREPARATION

Data preparation comprises of editing the data, coding the data, data capturing and data cleaning. These four components will be briefly explained.

5.9.1.1 Editing

Editing ensures that the questionnaire has been filled out properly and completely (McDaniel & Gates, 1998:351). According to Shao (1999:76), editing consists of checking completed questionnaires or other data collection forms for omission, incomplete or otherwise unusable responses, illegibility and obvious inconsistencies. The fieldworkers conducted field editing to check that the majority of questions were completed and that the handwriting of the respondents was legible. Central editing also took place as questionnaires were checked by the researcher and again when capturing the data to ensure that the information was correct, consistent with other information and complete. Emory and Cooper (1995:450) state that incomplete questionnaires will negatively affect the validity of the information collected. Incomplete questionnaires or non-responses can be handled in different ways. First, the respondent can be contacted again, but it was not possible in this study, since completion of the questionnaires were done anonymously. Secondly, the whole questionnaire can be disregarded as it is not useable (known as case-wise deletion). This method is appropriate if it is clear that the respondent either did not understand the survey or was not cooperating. A total of 15 incomplete and unusable questionnaires were eliminated from the study. Just the severe cases were dealt with

in this way, as applying this technique to all missing responses may lead to no valid cases in the data set, because each of them may have at least one missing data response in some variable (StatSoft, 1983-2004). The third option is to throw out only problem questions or the individual questions with non-responses and retain the balance of the questions. This method is known as pair-wise deletion and was used to deal with the remainder of the incomplete questionnaires, where only a few responses were missing. The last alternative is to replace the missing values with the mean value of that variable (Aaker, Kumar, Day & Lawley, 2005:382). This option was not used in this study, as substituting missing data with artificially created average data points may considerably change the values of correlations and decrease the variation of the scores (Statsoft, 1983-2004).

5.9.1.2 Coding

Coding is the assignment of numerals or other symbols to answers to enable the responses to be grouped into a limited number of classes or categories (Cooper & Schindler, 2003:456). Perreault and McCarthy (1996:114) point out that coding is the establishment of meaningful categories for responses collected by means of surveys or other data collection forms, so that the responses can be grouped into useable classifications. Codes can be assigned before or after a research study is completed. Jarboe (1999:78) defines pre-coding as the assignment of codes to the different responses on the questionnaire before the questionnaires are distributed. In this study, pre-coding was used as mainly closed-ended questions, multiple choice questions and scaled questions were used for which answers can be anticipated from the questionnaire. Pre-coding made the completion of data sheets unnecessary as the data was accessible directly from the questionnaire, saving time and money and decreasing the chances of coding errors.

5.9.1.3 Data capturing

Data entry or capturing is the task involved in the direct input of coded data into a software package that will ultimately allow the researcher analyst to manipulate and transform the raw data into useful information (Cant *et al.,* 2003:198). A personal computer was used as a data entry device. The data was entered directly from the

questionnaires. Data captured were entered into a MS Excel spreadsheet by data capturers of the University of Pretoria's Research Support Department. The data capturers checked for errors in the captured data. The researcher also checked the final captured data a second time to ensure that no wild codes or mistakes were made. The necessary changes were made by referring back to the original questionnaires.

5.9.1.4 Data cleaning

Data cleaning is an error checking process conducted after data entry and before data analysis to identify omissions, ambiguities and errors in the responses made during data entry (Diamantopoulos & Schlegelmilch, 2000:40). This could be prevented by not asking ambiguous questions and by pre-testing the questionnaire. Tustin *et al*. (2005:471) state that researchers have to attempt to clear the dataset of possible coding and data capturing errors. Data cleaning was done by making use of wild-code checks to detect codes that are not defined for a particular variable, as well as extreme case checks for responses to a variable that is far from ordinary. This study aimed to avoid mistakes during and immediately after the collection of data by working accurately and with numerous cross-checking.

After the data has been edited, coded, captured and cleaned, it can be analysed and described. The statistical tests used to analyse the data will be discussed in Section 5.9.2 and the results of the data analysis will be described in Chapter 6.

5.9.2 DATA ANALYSIS

Zikmund and D'Amico (2001:142) define analysis as the statistical and qualitative considerations of data gathered by research. The data needs to be analysed to extract the needed information to solve the research problem. In this study, the analysis of the data will be displayed in Chapter 6 by making use of descriptive statistics such as tables and figures, multivariate statistical techniques, hypothesis testing, and Cronbach alpha testing.

The results of the study will be presented using descriptive statistics by ways of frequencies, mean values and standard deviations. The mean measures the frequency

of occurrence and the measure of location, thus the average of the distribution of responses. The average is thus the sum of a set of values divided by their number. The mean is a measure of the central location of the distribution function of the variables and was used to determine the variance and standard deviation of the population. As standard deviation measures the spread of data about the mean, it implies that if the points are close to the mean, the standard deviation is small or low (indicating consistency or agreement) and if the points are far from the mean, the standard deviation is high (indicating inconsistency or difference). Tables and graphs will be used in Chapter 6 to display the descriptive statistics.

Statistical tests that make assumptions about the nature of the population from which the sample is drawn are known as parametric tests (Diamantopoulos & Schlegelmilch, 2000: 141). Data from this study can be classified as ordinal data (age), nominal data (gender, ethnic group and language) as well as interval data (sources of information and choice factors). Table 5.1 provides a summary of the levels of measurement and measurements scales used in this study.

Table 5.1: Measurement of variables

QUESTION NUMBER	VARIABLE	LEVEL OF MEASUREMENT	SCALE DESIGN
1.1 – 1.23	Higher education institution characteristics (choice factors)	Interval	Likert scale
2.1– 2.12	Source of information	Interval	Likert scale
2.13	Other sources of information	Nominal	Category
3	Gender	Nominal	Category
4	Age	Ordinal	Rank order
5	Ethnic background	Nominal	Category
6	Higher education institution attending	Nominal	Category
7	Home Language	Nominal	Category
8	Distance to higher education institution	Nominal	Category
9	Residence	Nominal	Category
10	Average grade	Nominal	Category

In this study, two tests namely MANOVA and t-test were used to test the hypotheses based on interval data. Firstly, the two-sample t-test was used to compare two groups on a variable measured at interval level. The null hypothesis tested by the two-sample t-test is that the two population means are equal, the alternative hypotheses being that the means are not equal (Diamantopoulos & Schlegelmilch, 2000).

Secondly, multi-analysis of variance (MANOVA) was used to test for differences in the mean values of several dependent variables (Lattin, Carroll & Green, 2003:389). The Wilks' lambda was the test statistic used to assess the overall significance of the MANOVA. Because the multivariate test of MANOVA shows only an overall significant difference, univariate analyses and a Scheffè post hoc test were performed to reveal more specific differences between groups on each of the identified choice factors.

Statistical tests that do not make use of stringent assumptions about the nature of the populations from which the sample data are drawn are known as non-parametric tests or distribution free tests. Cooper and Schindler (2001:495) explain that non-parametric tests are used to test hypotheses with nominal and ordinal data. Non-parametric tests were however not used in this study.

In this study, an MS Excel spreadsheet will be used to transfer the codes from the questionnaires onto the computer and the SAS software programme will be used to analyse the data. As previously mentioned, descriptive statistics and hypothesis testing will be done. For hypothesis testing, the significance level is set at 95 percent ($\alpha = 0.05$). According to Tustin *et al*. (2005:590), alpha (α) is denoted as the significance level and is used to indicate the risk that the researcher is willing to take in rejecting the true null hypothesis. A summary of the statistical tests used to test the stated hypotheses, as well as the objectives and appropriate questions from the questionnaire, are provided in Table 5.2.

Table 5.2: Objectives, hypotheses, questions and statistical tests

OBJECTIVE	HYPOTHESIS/RESEARCH QUESTION	QUESTIONS	TEST / STATS
Determine the relative importance of each of the choice factors that first year Economic and Management Sciences students used to select a higher education institution.	What is the relative importance of each of the choice factors that first year Economic and Management Sciences students use to select a higher education institution?	Question1: University characteristics considered (choice factors).	Descriptive statistics
Investigate the usefulness of the sources of information considered by first year Economic and Management Sciences students in the selection process as perceived by ethnic groups, gender groups and academic institutions attended.	What is the usefulness of the sources of information considered by first year Economic and Management Sciences students in the selection process as perceived by ethnic groups, gender groups and academic institutions attended?	Question 2: Usefulness of information sources.	Descriptive statistics
Determine whether students from different ethnic backgrounds differ regarding the importance they attach to choice factors when selecting a higher education institution.	H_1 Students from different ethic backgrounds differ regarding the importance they attach to choice factors	Question 5 and Question 1: Ethnic background and choice factors.	MANOVA
Determine whether students with different home languages differ regarding the importance they attach to choice factors when selecting a higher education institution.	H_2 Students speaking different home languages differ regarding the importance they attach to choice factors.	Question 7 and Question 1: Home language and choice factors.	MANOVA
Determine whether there are gender differences regarding the importance students attach to choice factors when selecting a higher education institution.	H_3 Male and female students differ regarding the importance they attach to choice factors.	Question 3 and Question 1: Gender and choice factors.	MANOVA
Determine whether students from different higher education institutions differ regarding the importance they attach to choice factors when selecting a higher education institution.	H_4 Students from different academic institutions differ regarding the importance they attach to choice factors.	Question 6 and Question 1: University attending and choice factors.	MANOVA
Determine if students that live further than seventy kilometres from a higher education institution make more use of campus visits or open days as a source of information than students living close-by.	H_5 Students that live further than seventy kilometres from the higher education institution attach more value to campus visits or open days as a source of information than students living close-by.	Question 8 and Question 2.4: Distance of parents' home from university and campus visits & open days.	Two-sample t-test
Determine if students that are resident in the province in which the higher education institution is located make more use of word-of-mouth as a source of information.	H_6 Students that are residents in the province in which the chosen higher education institution is located, value word-of-mouth (friends and other people) more as a source of information than students from other provinces.	Question 9 and Question 2.8: Residence in province and word-of-mouth.	Two-sample t-test
Determine if students with an average of seventy percent or more in Grade 12 make more use of higher education institutions' websites than students with a lower average in Grade 12.	H_7 Students with a Grade 12 average of seventy percent or more make more use of higher education institutions' websites than students with a lower Grade 12 average	Question 10 and Question 2.2: Grade 12 average and websites.	Two-sample t-test

5.10 SUMMARY

Chapter 5 introduced the empirical component of the study and set the scene for the discussions to follow in Chapter 6 (research findings) and Chapter 7 (recommendations and implications). Chapter 5 described the research process followed in this study. The research problem, research objectives and hypotheses were used to develop a research design. The research design delineated the structure of the investigation in such a way as to reach the objectives and solve the research problem. A sample of first year Economic and Management Science students from six higher education institutions were selected by making use of a non-probability convenience sample. A two-stage data collection process was followed: Firstly, a literature review followed by collecting quantitative data through a self-administrated questionnaire. Fieldwork was done by a designated fieldworker at each of the six institutions and collected over a three month period from February to April 2006. After data collection, the data was prepared by editing, coding, capturing and cleaning the data. The data was then analysed, making use of descriptive statistics and parametric tests. The analysis of the data will be presented in Chapter 6 and the findings in Chapter 7.