CHAPTER 8

RESEARCH METHODOLOGY AND DESIGN

8.1 INTRODUCTION

In the scientific approach to research, the researcher uses standardised methods for obtaining empirical answers to certain questions. Proper planning and preparation is the first critical requirement for any successful scientific research project. This should include the careful choice of a research strategy, demarcation of a population, specific sampling procedure and the use of appropriate statistical methods for data analysis. Theron (1992) emphasises the fact that suitable and proper research design, sampling methods and statistics ensure a soundly based, structured and systematic approach to scientific knowledge that can be checked for accuracy and the ability to generalise results to the population as a whole.

In this chapter the selection of an appropriate research design will be discussed. A brief description of the population will be followed by a discussion of the sample and the determination of a proper sample size that will be representative of the population in terms of all the independent variables discussed in Chapter 6. This representativeness will enable the researcher to generalise the research findings to the wider population. The selection of statistical methods for the analysis and a description of these methods will also be presented.

8.2 RESEARCH STRATEGY

The aim of this study, as discussed in Chapter 2, is to do a detailed analysis of work-related values, locus of control and leadership behaviour in a multicultural South African work force and their interrelations within the ambit of a transforming military organisation. The analysis will strongly focus on the effect of culture (specifically value differences) on transactional and transformational leadership behaviour displayed by leaders. The objective is to determine whether or not there are significant differences in leadership behaviours across cultures, in other words whether cultural differences (especially work-related values) prompt different leadership behaviours. The key question is whether or not the effectiveness of leadership behaviour is culture specific.

Value and leadership differences will be highlighted in terms of gender, age, home language, religion, level of education, occupational level, ethnicity and years of work experience as independent variables. The six value dimensions of Wollack et al (1971), the four value dimensions of Hofstede (1980), internality, and leadership styles will all be used as dependent variables. The achievement of the research aims in the study depends on obtaining information directly from the workforce about their work-related values and leadership behaviour. This information will be obtained from the sample subjects by posing questions (in the form of four questionnaires) about their personal preferences, intentions and behaviours. Questionnaires include two for work-related values, one for locus of control and one for leadership behaviour. Information regarding certain biographical variables will be obtained through the use of a separate questionnaire.

8.2.1 RESEARCH DESIGN

The research will be conducted by means of the survey method of data gathering. This method will be the most appropriate due to the researcher being able to visit all the various bases personally. All officers commanding were involved in making a sample of leaders on each base available for the survey. Although the survey method is the basic approach for this research, all the data could be considered as being part of an experiment, where multiple factor analysis of variance represents the main statistical method of data processing. Due to this, the research approach could also be described as a *posteriori* quasi-experimental design involving questionnaires. This design will be discussed in Section 8.2.2.

In order to achieve the study objectives as presented in Chapter 2, the statistical analysis of data will aim at achieving the following:

- Determining the construct and content validity of the four questionnaires to be used, viz:

  a. Internal Control Index (ICI) of Duttweiler (1984).

  b. Value Survey Module of Hofstede (1980).

  c. Survey of Work Values of Wollack et al (1971).

  d. Multifactor Leadership Questionnaire (MLQ) of Bass and Avolio (1997).

- Doing a reliability assessment of the four above-mentioned measuring instruments.

- Analysing the intercorrelations between work values, locus of control and leadership styles.

- Drawing a comparison by means of analysis of variance (in terms of work related values, locus of control and leadership styles) between four different ethnic groups.

- Drawing similar comparisons of the possible influence of age, language, religion, level of education, occupational level and years of work experience.

- Doing a discriminant analysis of leadership styles in terms of work related values and locus of control.

- Evaluating the appropriateness and suitability of a transformational leadership approach across all cultural groups in South Africa

8.2.2 POST HOC (A POSTERIORI) QUASI-EXPERIMENTAL DESIGN

As opposed to planned (or *a priori*) comparison, Shavelson (1981: 469) refers to post hoc comparison as a comparison of means which has not been planned but which, on the basis of the sample data, looks interesting to the researcher and allows him to find out "…where the differences occurred which gave rise to the significant, overall *F*[1]". He states that, if the overall *F* is significant, at least one out of all possible comparisons between pairs of means will be significant. Although statistically less powerful, post hoc comparisons are useful and are very often used in social science research. One of the most well-known post hoc comparison tests is the Scheffé test (Bohrnstedt et al, 1982). They define post hoc comparison as a "hypothesis testing of the differences among population means carried out following an analysis of variance". The notion of a contrast is used to make multiple comparisons among a set of means. A contrast is a set of weighted population means that sum to zero, used in making post hoc comparisons of treated groups. It is usually referred to with the label Ψ (psi). The basic requirement for using post hoc comparisons is that the overall F in the analysis of variance must be significant.

A quasi-experimental design differs from true experiments in that it lacks the random assignment of subjects to an experimental and a control group (Babbie, 1989). It is a research plan that has some but not all the validity features of an experimental design. Manipulations of the independent variable are quite difficult and, under certain circumstances, even impossible (Dooley, 1990:198). The emphasis in quasi-experimental designs is whether an independent variable is an indicator of whatever the real cause may be and not the actual cause of the dependent variable (Dane, 1990: 105). A quasi-

[1] The F-ratio is a test statistic formed by the ratio of two mean-square estimates of the population error variance (Shavelson, 1981: 469).

experimental design is explained by Mason et al (1989: 127) as an alternative to experimental design in that it can be carried out in field settings and does not need to comply with the requirements of equalisation of groups by means of the random assignment of subjects.

Quasi-experiments should be employed in research settings where the basic elements of a true experiment cannot be set up (Baker, 1988: 223). In the research at hand the researcher wishes to determine the effect of the independent variable(s) on the dependent variable(s) and also the influence of nuisance variables. Because the study will be carried out in the natural setting where the experimental event(s) occur, the researcher may be forced to use a quasi-experimental design. In these natural settings, where the researcher does not have total control, quasi-experimental techniques are employed to deal with the threats of internal and external validity (Mason et al, 1989: 127). Dooley (1990: 183) refers to internal validity as the truthfulness of the claim of a causal linkage between variables internal to the design, while external validity is referred to as the extent to which research findings may be applicable to other populations, other times and other settings.

A disadvantage of the quasi-experimental design is its susceptibility to the threats of statistical regression, history, maturation, testing and instrumentation. These are all sources of internal validity that the researcher has to take into account when planning the research design. Babbie (1989: 221) refers to the history effect as the influence of those events that may occur during the course of the experiment that will confound the experimental results. Maturation is the result of continuous growth and change in people. Even in shorter experiments these changes may affect the results of experiments. Instrumentation effects refer to changes in the manner in which the dependent variable is measured (i.e. the use of different questionnaires to measure the same dependent variable). Chadwick, Bahr and Albrecht (1984: 178) explain (statistical) regression as the tendency for extreme behaviour to be replaced by less dramatic behaviour. When subjects start out with extreme scores on the dependent variable (i.e. extremely low) there is the inherent possibility that the scores of these subjects can only stay the same or increase. Babbie (1989: 223) warns that the danger in this is that changes occurring by virtue of subjects starting out in extreme positions will be attributed erroneously to the effects of the experimental stimulus.

Depending on the research setting the researcher can choose one of several quasi-experimental designs. Mason et al (1989: 129-137), Baker (1988: 223-225) and Howard (1985: 117-129) present six such designs, viz simulated before-after design, non-equivalent control group design, regression discontinuity experiments, time-series experiments, counterbalanced design and equivalent-time-samples design. In the simulated before-after design a group of subjects are identified, all of whom will be exposed to an intervention. The group is then randomly divided into two parts, of which one part will be pretested, but not posttested. The other half will be posttested but not pretested (the experimental group). The advantage of this design is that the two groups are equivalent at the time of the pretest (Howard, 1985: 121). It is, however, open to the effects of history and maturation as described above. Non-equivalent control group designs are used where random assignment to groups is not feasible (Baker, 1988: 223). Both the experimental and the control group take a pretest as well as a posttest. In this design only the experimental group is exposed to the experimental variable and is then compared to a similar (not randomly selected) control group that was not exposed to the experimental variable. The design could be presented as follows (Mason et al, 1989: 129):

O*(1)* ***X*** ***O****(2)* ➔ *(experimental group)*
O*(1)* ***O****(2)* ➔ *(control group)*

X = exposure to experimental variable (treatment)

The regression-discontinuity design is used in cases where it would not be practical to have another group exposed to the treatment or to serve as a control group. The design indicates differences that occur at the point of treatment which would differentiate post-treatment scores of those having been treated from those of the group not receiving the treatment. Cook & Campbell (1979: 137) mention that the design is especially appropriate "when people or groups are given rewards or those in special need are given extra help and one would like to discover the consequences of such provisions." Time-series designs (of which there are two types, viz interrupted and multiple time-series) generally use a large set of already collected data which indicate rates over standard intervals of time. Some other event (treated as the independent variable) is then superimposed on this time line data to determine whether there is a change at the point where the event occurred. The dependent variable is measured several times before and after the introduction of the independent variable. In a counterbalanced design, there are several different treatments and several respondents and each respondent is presented with each treatment condition in random order. The design could be explained by the matrix in Figure 8.1:

Figure 8.1: A counterbalanced quasi-experimantal design.

Group or Subject	Time or setting			
	1	2	3	4
A	X_1O	X_4O	X_3O	X_2O
B	X_2O	X_3O	X_4O	X_1O
C	X_3O	X_1O	X_2O	X_4O
D	X_4O	X_2O	X_1O	X_3O

(X_1O) = treatment condition 1
(X_2O) = treatment condition 2
(X_3O) = treatment condition 3
(X_4O) = treatment condition 4

(Source: Mason et al, 1989: 132)

The equivalent-time-series design involves an experimental setting where each subject serves repeatedly under the experimental and control conditions of an experiment. It could be presented as follows (Mason et al, 1989: 134):

$X_1O \quad X_0O \quad X_1O \quad X_0O$, etc

where,

X_1O represents the application of the experimental variable and X_0O represents the control condition or some alternate condition.

For a more effective version of the design the pattern of X_0O and X_1O may also be randomised instead of alternated.

The present study involves a single group of subjects that are measured on a number of dependent variables and may therefore be referred to as a one-shot case study (as described by Theron, 1992: 336). It could be represented by the following formula :

X -----------------O

where X = exposure to the experimental variable, and
O = observation of the group (measurement).

The survey research method (discussed below) will be used using an *ex post facto* design. In *ex post facto* design the occurrence of the event(s) under study has already taken place, i.e. the researcher enters the situation after the event (Mark, 1996: 166). Additional variables are introduced into the data analysis to determine their effect on the observed relationship between the independent and dependent variables under study. Evidence is gathered to support or reject a hypothesis. In this study only one measurement will be taken with the four scales described earlier to determine the interaction between the independent and dependent variables. Although the independent variables have a controlling effect, there will be no control group in the research setting.

8.2.3 SURVEY RESEARCH

Today, according to Babbie (1989: 236) survey research is probably the most frequently used mode of observation in the social sciences. It is, for example, the most common method reported in recent articles of the *American Sociological Review*. It could be regarded as the best method available to the scientist interested in collecting original data about a population too large to observe directly. Typically, the researcher selects a sample of respondents from a certain population and administers a standardised questionnaire(s) to them.

Baker (1988: 472) defines survey research as "a research method that analyses the responses of a defined sample to a set of questions measuring attitudes and behaviours … a method of collecting data in which a specifically defined group of individuals are asked to answer a number of identical questions". Dane (1990: 338) defines survey research as "a method of obtaining information directly from a group of individuals", while Mason et al (1989:52) view it as "… a technique to study the distribution of characteristics in a population". Chadwick et al (1984: 442) see survey research as "a research technique that puts questions to a sample of respondents by means of a questionnaire or an interview". Most of these definitions seem to emphasise the fact that data is collected from a portion (sample) of the population in order to obtain characteristic information about the population as a whole. The sample size, which in survey research is generally large, distinguishes it from other research strategies and methods. Babbie (1989: 338-253) discusses three different methods of survey research, viz self-administered questionnaires, interview surveys and telephone surveys, and then notes that, if complete anonymity is offered, self-administered surveys are more appropriate in dealing with especially sensitive issues like controversial or deviant attitudes or behaviours. According to Theron (1992: 337) random assignment, manipulation of the independent variable and testing of the cause-effect hypothesis seldom form part of the survey research.

Apart from the fact that survey research may test specific hypotheses, it also has the aim of describing the characteristics of a select sample and evaluating the presence and effects of various factors (Baker, 1988: 16). The process starts with the selection of an appropriate and valid measurement. A valid measurement is regarded as a questionnaire with questions that measure the concept(s) that the researcher intends. This calls for questions to be worded carefully and unambiguously so that the gap between what the researcher wants to measure and the actual results of the survey is as narrow as possible. Therefore, the designing of questions is a critical phase of the survey.

After selecting either the complete questionnaire or the appropriate items for inclusion in the questionnaire to measure those concepts the researcher wants to measure, the researcher has to decide on the modes of eliciting information from the respondents and the modes of returning information (Baker, 1988: 168). The primary modes of eliciting information that are both based on a fixed set of questions, are the completion of questionnaires and conducting face-to-face or telephone interviews. In this case a set of four scales (as discussed earlier) were selected for inclusion in the questionnaire. The next step was to select the respondents for participation (sampling). In selecting respondents for the survey, an important criterion is that the questions should apply to the population from which the respondents are selected. The population in this research consists of leaders at all levels of one of the arms of service of the SANDF. The sampling process will be discussed under section 8.4.

8.2.4 DESIGN AND CONSTRUCTION OF QUESTIONNAIRES

The survey researcher may make use of four different types of questions to be included in the questionnaire, each with a specific purpose. Baker (1988: 173-174) lists them as closed-ended questions, open-ended questions, contingency questions and matrix questions. Matrix questions give the respondent the opportunity to answer sets of questions with similar questions, for example the response categories of a Likert scale[2]. Typically the respondent is asked to either "Strongly Agree", "Agree", "Have no opinion", "Disagree", or "Stronly Disagree" with a set of similar questions or statements. He then selects one of these responses for each question. The Work Value Survey of Wollack et al (1971), the Value Survey Module of Hofstede (1980), the Internal Control Index (ICI) of Duttweiler (1984) and the Multifactor Leadership Questionnaire (MLQ) of Bass et al (1997) as they are used in this research, are all examples of Likert scales.

Open-ended questions allow for more detailed answers to questionnaire items. Baker (1988: 174) points out that most of the guidelines for constructing open-ended questions are focussed on ensuring that the respondent does not skip a question, especially due to the fact that forced choice questions and matrix questions are more likely to be completed by respondents than open-ended questions. He suggests that a specific number of lines be left available to ensure a more precise response. On the other hand, too many lines may

[2] A type of composite measure developed by Rensis Likert in an attempt to improve the levels of measurement in social research through the use of standardised response categories in survey questionnaires (Babbie, 1989: G5).

cause the respondent to skip the item. Some other guidelines include putting interesting questions first (to encourage the respondents to complete the questionnaire) and putting sensitive questions near the end of the questionnaire. Questions should also be worded in such a way that respondents understand them. The disadvantage of written (open-ended) responses is that they require much more time and thought from the respondent to answer them. These questions are also much more difficult to code. In the present research only the five introductory (leadership-related) questions to the MLQ are in an open-ended format.

8.2.5 INSTRUMENTS INCLUDED IN THE SURVEY QUESTIONNAIRE

At this point a brief reference to the questionnaires used is appropriate. The researcher, as was mentioned earlier, decided on using two work-related-value scales, one locus of control scale and one leadership questionnaire. Biographic information was obtained by means of a short separate questionnaire.

8.2.5.1 The Internal Control Index (ICI) of Duttweiler (1984)

The test items of the ICI are based on those variables that proved to have the most pertinent relation with internal locus of control, namely cognitive processing, autonomy, resistance to influencing, delaying of reward and self confidence. The 28 items of the ICI are assessed on a five-point scale. A reported reliability coefficient of 0.84 was obtained for the test (Duttweiler, 1984). De Kock (1995) reported a 0.767 reliability coefficient.

8.2.5.2 Evaluation of work-related values

As indicated in Chapter 3, a two-dimensional approach as well as a multi-dimensional approach will be followed in the analysis of work-related values. The Survey of Work Values of Wollack et al (1971) divides the Protestant Ethic in intrinsic and extrinsic aspects of work and will be used as a two-dimensional instrument. The Value Survey Module of Hofstede (1980) provides a multi-dimensional approach to the analysis of work-related values and will be applied as the second instrument in the study.

8.2.5.2.1 The Survey of Work Values of Wollack et al (1971)

8.2.5.2.1.1 Purpose of the scale

The Survey of Work Values is based on the Protestant Work Ethic and was constructed to measure a person's attitude towards work in general rather than his feelings about a specific job. It measures several areas of work-related values. Wollack et al (1971) view the concept of work-related values as the meaning an individual assigns to his work.

8.2.5.2.1.2 Composition of the scale

Because the intrinsic aspects of work (i.e. work as rewarding in itself) form such an important part of the Protestant Ethic, Wollack et al (1971) selected three dimensions of the Protestant Ethic comprising the intrinsic aspects of work viz “pride in work”, “job involvement” and “activity preference”. The extrinsic aspects of work were also addressed by including the following sub-scales: “attitude towards earnings” and “social status on the job”. Two further dimensions which were regarded to be of a mixed character were included, viz “upward striving” and “responsibility towards work”. After determination of internal validity of the scale through a principal components analysis with quartimax rotation (see Chapter 7), the sub-scales were reduced to six, each with nine items ("responsibility towards work" was eliminated).

8.2.5.2.2 The Value Survey Module of Hofstede (1980)

In his research conducted on work-related values in 40 different international cultures, Hofstede (1980) found four dimensions of difference between national cultures viz. “power distance”, “individualism vs collectivism”, “uncertainty avoidance” and “masculinity vs femininity”. These four value dimensions were empirically determined by means of a factor analysis of mean scores of respondent samples from different countries.

A questionnaire on work-values containing 120 questions was initially administered to a sample of employees in 40 countries (Hofstede, 1980: 68). A five-point Likert scale was used to evaluate responses. After analysing the data by means of analysis of variance and factor analysis, Hofstede further refined the instrument, which he then named the Value Survey Module.

A discussion of validity and reliability of the Value Survey Module was provided in Chapter 7 and will therefore not be repeated here.

8.2.5.3 Evaluation of leadership behaviour

8.2.5.3.1 The Multi-factor Leadership Questionnaire (MLQ) of Bass et al (1997)

The Multifactor Leadership Questionnaire (MLQ) consists of two different questionnaire forms, each consisting of 45 items: the Self-Rating Form, used by superiors/supervisors themselves as leaders and the Rater Form used by followers to rate their leaders. These followers can represent three different organisational levels, viz above their ratee, same level as ratee, or below their ratee. A computer generated report, the MLQ Report, provides information on how raters perceive the behaviours of their leaders along a full range of leadership styles (as discussed in Chapter 5).

8.2.5.3.1.1 Advantages of using the MLQ

For Bass and Avolio (1997: 3) the MLQ represents an effort to capture a broader range of leadership behaviours, from Laissez Faire to idealised leadership, while also differentiating ineffective from effective leaders. These authors also provide a number of advantages for using the MLQ in evaluating leadership behaviour:

- The range of ineffective and effective leadership behaviours in the MLQ is typically much broader than other leadership surveys commonly in use.

- The MLQ is suitable for administration at all levels of organisations and across different types of production, service and military organisations.

- The MLQ has 360 degree capabilities and can be used to assess perceptions of leadership effectiveness of leaders from many different levels of an organisation.

- The MLQ emphasises development. It includes items that measure a leader's effect on both the personal and intellectual development of self and others.

- The MLQ is based on a model that is easy to understand.

8.2.5.3.1.2 Factor structure of the MLQ

The MLQ has undergone numerous revisions and is now available as an instrument containing 36 leadership items and nine leadership outcome items (Bass et al, 1997). In a convergent and discriminant validation study containing 3 570 cases, the choice of the 45 items in the MLQ as the best indicators of their constructs was confirmed by Avolio, Bass & Jung (1996).

8.2.6 ADMINISTERING THE QUESTIONNAIRE

The next step after questionnaire construction, is the administering of the questionnaire. The two main strategies for collecting data through self-administered questionnaires are explained by Chadwick et al (1984: 147). They may either be delivered to individual respondents and collected after a few days or they may be administered directly to groups. Chadwich et al (1984:147) mention that the latter is much more efficient. In this way the data collection is both easier and quicker. In addition, the researcher also has the opportunity to explain the purpose of the research, to highlight the instructions for completion and to immediately handle queries and uncertainties.

The questionnaires in this survey were not distributed into the research field. Instead, the questionnaires were administered to a group of randomly selected leaders at each SAAF unit. Each unit of the organisation (these units are spread throughout the country and consist of bases, squadrons and depots or servicing units) was visited, during which groups of between 40 and 70 respondents were gathered in a lecture room for voluntary completion of the questionnaires. Members not wanting to participate were allowed to withdraw. By using this approach the following were ensured:

- The attention of respondents could be specifically focussed on all important instructions.

- Any questions could be dealt with immediately.

- Questionnaires were received back immediately. A considerable amount of time could thus be saved. The problem regarding questionnaires not being received back was almost eliminated.

The survey was done anonymously and participants were requested not to indicate any names or personal force numbers on the answer sheets. To encourage honest participation, respondents had the opportunity to indicate an eight-digit "secret code" so as to allow for personal feedback to be given to anyone interested. On completion the respondents handed back the questionnaires to the researcher.

8.3 THE POPULATION

Babbie (1989: 169) defines the term population as "the theoretically specified aggregation of study elements". Because it is not possible to guarantee that every element meeting the theoretical definition of the population actually has a chance of being selected in the sample, he distinguishes the term "study population" from the term population. A study population is "that aggregation of elements from which the sample is actually selected" (Babbie, 1989: 170). De la Rey (1978: 16) offers a wider definition of a population: "all the species, persons, or objects being present at a certain place and time holding a specific characteristic". He emphasises that, in order to satisfy the demands of scientific verification, the researcher should demarcate and define the population as precisely as possible. The population to which the study is directed consists of all the so-called "uniformed" or military leaders of the SAAF. Leaders in this case are defined as all non-commissioned officers holding a rank of sergeant or higher, all warrant officers as well as all officers (excluding candidate officers) having followers reporting directly to them. The composition of the population in terms of members per rank group is reflected in Table 8.1. The organisation employs approximately 12 000 members throughout the country which include all Permanent Force, short term and civilian employees. As the civilian members will not be included in the study, the actual size of the organisation's military workforce is 9162. The survey population of military leaders, as per definition, totals 6781. This population consists of 1104 female members and 5677 male members. A large number of occupational musterings are involved.

Furthermore, a large variety of ethnic groups are represented within the organisation, although the largest percentage still constitutes white employees. The SAAF is a typical large public service organisation where the ages of employees vary between 17 and 60. Levels of seniority are determined through a fixed military rank structure (from airman to lieutenant general).

Table 8.1: Composition of SAAF workforce per rank.

Rank	Female	Male	Total	Leader Total
Lieutenant General	0	1	1	**1**
Major General	0	7	7	**7**
Brigadier General	2	33	35	**35**
Colonel	11	164	175	**175**
Lieutenant Colonel	53	374	427	**427**
Major	78	308	386	**386**
Captain	74	291	365	**365**
Lieutenant	68	214	282	**282**
2nd Lieutenant	6	23	29	**29**
Warrant Officer I	55	594	649	**649**
Warrant Officer II	96	653	749	**749**
Flight Sergeant	282	1223	1505	**1505**
Sergeant	379	1792	2171	**2171**
Corporal	158	1407	1565	
L Corporal	22	223	245	
Airman	93	471	564	
Pioneer	0	1	1	
Senior Pioneer	0	1	1	
Chief Pioneer	0	5	5	
Totals	1377	7785	**9162**	**6781**

8.4 SAMPLING PROCEDURE

Baker (1988: 144) defines a sample as “a selected set of elements (or units) drawn from a larger whole of all the elements, the population”. The researcher has a choice of many sampling methods and his most important concern is to ensure that the sample is representative of the wider population in terms of the variables studied. A representative sample can only be guaranteed by drawing a sample structurally and methodically, thus enabling the researcher to obtain reliable results (De la Rey, 1978: 16). If the researcher wishes to generalise the questionnaire responses to a wider population, he has to develop a probability sample. Baker (1988: 469) defines probability sampling as “a sample designed according to the rules of probability, which allows a determination of how likely the members of the sample are to be representative of the population from which they were drawn”. Without a probability design findings cannot be generalised. In fact, a number of statistical tests assume that the data being used have been collected according to the rules of probability. Baker (1988: 155) warns that these tests will be meaningless if they are applied to findings from a nonprobability sample.

In this research a sample of 509 members were drawn from the population described in the previous section. It could be regarded as a stratified sample[3] where the different units of the organisation form the various strata. These strata are homogeneous with regard to variables like gender and age, but also rather heterogeneous with regard to the distribution of population groups. Due to the low percentage of blacks being employed in the organisation (and an even lower percentage of these members being in leadership positions) the researcher selected extra black members as far as possible in an attempt to increase the percentage of blacks in the sample. The reasons for the use of a stratified sampling method are as follows:

- Because the organisation's employees have a wide geographical distribution, the researcher was compelled to visit the various regions. It would have been impossible to take a representative sample at one central point.

- Visits to various units would enable personal contact with respondents.

- Certain groups of employees are represented stronger in some strata than others. There are for example, more coloureds located at the Cape units than at other units. For this reason the researcher ensured a “coloured-heavy” sample in the cape area and an “Indian-heavy” sample in the Durban area. Similarly, the researcher had to ensure the inclusion of as many as possible senior leaders (Colonel to Lieutenant General) from the Headquarters due to the non-availability of an adequate number of these members at the other bases. In this way it was ensured that the total sample is representative with regard to all biographical variables.

The only disadvantage of the chosen sample is that the number of respondents in a stratum is not exactly proportional to the total number of people in each stratum of the population. However, this will not have any negative effect on the study, as no statistical processing of data per stratum will be done. The sample of 509 members proves to be representative in terms of gender, age, educational level, seniority and population group (ethnicity).

[3] A stratified sample is a type of random sample in which the researcher first identifies a set of mutually exclusive categories and then uses a random selection method to select cases (respondents) in each case (Theron, 1992).

8.5 STATISTICAL METHODS IN DATA PROCESSING

8.5.1 INTRODUCTION

As was indicated in chapter 2, the researcher wants to examine the nature of locus of control and work-related values, the intercorrelation between these two constructs and the effects thereof on the behaviour of transformational leaders. The researcher hopes to ascertain the existence of significant differences in terms of work-related values and leadership preferences amongst different population groups and also to ascertain the influence of the independent variables i.e. gender, age, religion, seniority and population group. The data collected as described earlier in this chapter, will be extensively analysed by statistical tools as described by Mark (1996), Tabachnick and Fidell (1983), Ferguson (1981), Rowntree (1981), Theron (1992), Huysamen (1991), Bohrnstedt et al (1988), and De la Rey (1978). The major tools of statistical data analyses will be descriptive statistics, analysis of variance, discriminant analysis, correlation statistics (i.e. Bravais-Pearson), and non-parametric statistics. Non-parametric statistics used, will include Spearman's rho, Kendall's Tau and Kriskal-Wallis one-way analysis of variance. Multiple regression will be used to determine how the first five leadership questions predict leadership behaviour.

8.5.2 DESCRIPTIVE STATISTICS

Descriptive statistics can be described as the statistics used to summarise data (Mason et al, 1989: 428). It provides a description of the features of a set of observations, viz percentage, modes, means, frequency distribution, kurtosis, skewness, variance, the standard error of the mean, and standard deviations (Bohrnstedt et al, 1988: 492). Descriptive statistics, according to Cooper and Schindler (1998:427), could be divided into measures of location, measures of spread, and cross tabulations. For nominal data, each category is represented by its own numerical code, while ordinal data are ordered in hierarchical form, varying from lowest to highest.

Bohrnstedt et al (1988: 491, 496) describe the normal distribution as the most important and most significant distribution. It is a smooth, bell-shaped theoretical probability distribution for continuous variables[4] that can be generated from a formula. Distribution is described by the characteristics of location, spread and shape. Cooper et al (1998: 427) list the following characteristics of descriptive statistics:

[4] A variable that in theory can take on all possible numerical values in a given interval.

- The shape of a distribution is just as consequential as its location and spread.
- Visual representations are superior to numerical ones for discovering a distribution's shape.
- The choice of summary statistics to describe a single variable is contingent on the appropriateness of these statistics for the shape of the distribution.

8.5.2.1 Measures of central tendency

8.5.2.1.1 The Mean

Measures of central tendency include the mean, mode and the median. The mean is the most frequently used statistic for both interval and interval-ratio data (Cooper et al, 1998: 428) and is described as the arithmetic average, which is symbolised by $\overline{X}$ (Bohrnstedt et al, 1988). In the case of the distribution containing extreme scores, the mean can be misleading. Cooper et al (1998) offer the following formula for calculating the mean:

$$\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

8.5.2.1.2 The Mode

The mode is a further measure of central tendency. It refers to the most frequently occurring value in situations where different values of X occur more than once. A modal value can therefore not be calculated when all values of X occur with equal frequency and where the frequency may be equal to or greater than one. The mode is a point as reference and, together with the mean and median, may be used for analysing spread and shape (Ferguson, 1981:56).

8.5.2.1.3 The Median

The median is the mid point of distribution and divides an ordered frequency distribution into two equal halves. One half of the distribution falls above and the other below the median (Bohrnstedt et al, 1988). Due to the fact that the median has resistance to extreme scores, it is a preferred measure of interval-ratio data. In cases where even numbers of observations occur in the distribution, the average of the two middle scores represents the median.

8.5.2.2 Measures of variation

The measures of variation that are to be calculated are skewness, kurtosis, variance, the standard error of the mean and the standard deviation. Ferguson (1981: 40) refers to skewness as the dispersion of a distribution based on the observation that “when a distribution is symmetrical, the sum of the cubes of deviation above the mean will balance the sum of cubes below the mean”. When a distribution is skewed to the right, the sum of cubes of deviations above the mean will be higher than the some of those below the mean and *vice versa*.

Kurtosis indicates a distribution’s peakedness or flatness. In distributions with a peaked or leptokurtic shape, the scores cluster or pile up in the center. The scores of platikurtic (flat) distributions are evenly distributed. A normal kurtosis has a value of 0.263. The kurtosis value of a peaked (leptokurtic) distribution is greater than 0.263, while the kurtosis value of a flat distribution is less than 0.263. Cooper et al (1998:430) provide the formula for kurtosis as follows:

$$KU = \frac{M4}{M_2^2} - 3 = \frac{\sum x4 / N}{(\sum x^2 / N)^2} - 3$$

The variance is the average of the squared deviation scores from the distribution’s mean, and is therefore a measure of score dispersion about the mean. In cases where all the scores are identical, the variance is 0. A greater variance is an indication of a greater dispersion of the scores. S^2 is used as the symbol for the sample variance and the Greek letter sigma (σ) for the population variance. The formula for S^2 is:

$$S^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{n-1}$$

(Cooper et al, 1998:429)

The variation (σ^2) is always positive and is called the sum of squares ($\sum x^2$).

8.5.2.2.1 Standard deviation

The standard deviation is the square root of the variance. It is also used to describe a dispersion of a distribution. Du Toit's (1963:37) formula is as follows:

$$S = \sqrt{\frac{\sum x^2}{N-1}}$$

According to Theron (1992:370) the standard deviation is a measure of the average of the scores' deviations of the mean. In a normal distribution, two-thirds of the observations lie within one standard deviation of the mean.

8.5.2.2.2 The standard error of the mean

Theron (1992:370) describes the standard error of the mean as "the standard deviation of sample means in a sampling distribution". A greater variability among sample means indicates a greater chance of incorrect inferences about the population mean from a single sample mean. It provides the researcher with information about the amount of error that is likely to be made in the process of inferring the population mean from the sample mean (Shavelson, 1981:305).

8.5.2.3 Frequency tables

Howell (1999: 28) describes a frequency distribution as a distribution that plots the values of the dependent variable against their frequency of occurrence, i.e. the number of times each value of the variable is observed in the sample. Frequency tables, therefore, consist of information about the values of variables (Theron, 1992:371). In tables, percentages and cumulative percentages are used to describe the sample.

8.5.2.4 Cross-tabulation

Bohrnstedt et al (1988:101) describe a cross-tabulation as a "tabular display of the joint frequency distribution of two discrete variables which has r rows and c columns". It therefore indicates the joint outcome of two variables. Such a table can be used to determine whether two variables are in fact related as hypothesised.

8.5.3 CORRELATION STATISTICS

Bohrnstedt et al (1988: 491) define the correlation coefficient as "…a measure of association between two continuous variables that estimate the direction and strength of linear relationship". Known as the Bravais-Pearson product-moment correlation coefficient, it is symbolised by r_{xy}. The two variables should be measured on either an interval or a ratio scale. Due to the correlation coefficient also indicating the strength of a relationship, it varies over a range of +1 to –1. The sign signifies the direction of relationship (Cooper et al, 1998:517). A value of –1 represents a perfect inverse association, while a value of +1 refers to a perfect positive correlation. A zero indicates that there is no relationship at all. A stronger correlation therefore indicates that y is better predicted by x.

8.5.4 ANALYSIS OF VARIANCE

Ott et al (1990:695) define analysis of variance (ANOVA) as "a procedure for comparing more than two populations", while Bohrnstedt et al (1988:219) view ANOVA as a statistical method to test the hypothesis that "…the sample means of two or more groups come from the same rather than different populations". ANOVA could be seen as a method to determine whether or not differences between groups exist (Theron, 1992:343).

Theron (1992:343) notes that it is also possible to test the strength of association between independent and dependent variables, for which a variety of techniques are available. The essential question in ANOVA is how much of the total variance in the dependent variable can be explained by the independent variables and how much is left unexplained.

Bohrnstedt et al (1988:222) advance the following formula for the general ANOVA model with one independent variable (IV):

$$Y_{ij} = \mu + a_j + e_{ij}$$

where,

e_{ij} equals the difference between an observed score and the score predicted by the model (error term).

The formula indicates that the score observation (*i*), which is a member of group *j* (therefore Y_{ij}), is a function of a group effect, (a_j), plus the population mean (*y*) and random error (e_{ij}). The error term is needed to take into account that not all observations in the subgroup *j* has the same Y_{ij}.

One-way variance analysis allows the researcher to measure the effect of an independent variable (IV) on a dependent variable (DV) (Theron, 1992:345). In factorial ANOVA (another technique of variance analysis), two IV's are simultaneously investigated. This technique involves two bases of classification, which are called factors.

In a two-way factorial ANOVA, the sum of squares is divided into three parts, namely a "between-rows" sum of squares, a "between-columns" sum of squares, and an "interaction" sum of squares (Ferguson, 1981:253). The total sum of squares of all the observations about the grand mean is presented as follows:

$$\sum_{r=1}^{R} \sum_{c=1}^{C} \sum_{i=1}^{n} (X_{rci} - \overline{X}...)^2$$

ANOVA, being analogous to the levels test, the parallelism test and the flatness test, allows for analysis of variance to be used for conducting a profile analysis. Here, treatments correspond to rows and dependent variables to columns. (Harris, 1975:81).

8.5.5 DISCRIMINANT ANALYSIS

Cooper et al (1998:525) classify discriminant analysis as a dependency technique. It is used for the classification of people or objects into (two or more) groups in order to establish a procedure for the finding of the predictors that best classify subjects. Discriminant analysis can also be used to analyse known groups for determining the relative influence of certain factors.

Discriminant analysis can serve as a measure for doing profile analysis. Profile analysis is viewed as a generic term of all methods concerning groupings of persons" (Nunnally, 1967:372). He mentions two purposes of the analysis. Firstly, it distinguishes groups from one another on the basis of scores in a data matrix. Secondly, it is used to assign individuals to groups in terms of the profile score. In the present study group membership is known and the purpose of the discriminant analysis is to distinguish the various groups on the basis of scores in the data matrix.

Pretorius (2004:155) describes discriminant analysis as MANOVA turned around. Due to the fact that MANOVA can be used to determine whether group membership produces reliable differences on a combination of dependent variables, the discriminant procedure can be applied when using a combination of variables to predict group membership. In this procedure the IV's are predictors and the DV's are the groups (Tabachnick et al, 1989:506).

In the present study the discriminant function analysis is used for clustering profiles. This analysis is employed in cases where groups are defined *a priory*. Here the purpose is to distinguish the different groups from one another based on scores obtained in a series of tests (Nunnally, 1967:388). Theron (1992:355) warns that discriminant function analysis is sensitive to multivariable outliers[5].

8.5.6 STUDENT's T-TEST

The Students t-test is an inferential statistic used by the researcher to decide whether observed differences between two sample means arose by chance, or represents a true difference between populations, i.e. whether or not to reject the null hypothesis of no difference between the means of the two groups (Shavelson, 1981:419). As the decision cannot be made with complete certainty, the researcher has to determine the probability of observing the difference between the sample means of the two groups under the assumption that the null hypothesis is true. Bohrnstedt et al (1988:204-205) advance the formula for a test to determine the probability of observed sample means occurring in the population:

$$S^2 = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}$$

where,

$N_1 + N_2 - 2$ are the degrees of freedom which are associated with S^2.

De la Rey (1978:71) lists certain assumptions, which have to be met prior to the t-test being used:

- The scores in the respective populations must show a normal distribution.
- The t-test, being based on sample means, requires the two samples to be big and of equal or almost equal size.
- The measurements must be on either interval or ratio level.
- The scores in the groups should be randomly sampled from their populations.

[5] Outliers are defined as "cases with extreme values on a variable or combination of variables, which unduly influences the averages of scores and invalidates the generalisability of the solution to the population" (Theron, 1992:355).

8.5.7 NON-PARAMETRIC STATISTICS

The two non-parametric statistics that will be discussed here are the Kruskal-Wallis one-way analysis of variance (for three or more independent samples) and the Mann-Whitney U-test (for two independent samples). De la Rey (1978:113) states that one or more of certain assumptions needs to be met for the application of non-parametric statistics:

- the scores distribution has to be skewed;
- measurements must be on either nominal or ordinal level;
- the sample size must be small (N=$\leq$30);
- it is a situation in which it is impossible to make certain assumptions in regard to the sample; and
- it is impossible to realise certain research aims due to appropriate parametric statistics not being available.

8.5.7.1 Mann-Whitney U-test

The Mann-Whitney U-test, a distribution-free non-parametric test, is used for comparing the central tendency of two independent samples. The test may also be applied to normally distributed populations. It serves as an alternative to the t-test, but without the t-test's limiting assumptions (Theron, 1992:366). The Mann-Whitney U-test is based on the ranking of scores. This ranking is a sophisticated mathematical operation, which can be performed on ordinal level data. Siegel (1956:120) presents the following formula to compute U:

$$U = N_1 N_2 + \frac{N_1(N_1+1)}{2} - \sum R_1$$

where $\sum R_1$ = the sum of the ranks for sample 1, whose size is N1.

When determining the value of U, the researcher has to conduct a test of significance. In doing so a z-score is obtained with the formula:

$$Z(obtained) = \frac{U - \mu_u}{\sigma_u}$$

where U = the sample statistic,

μ_u = the mean of the sample distribution of sample Us; and

σ_u = the standard deviation of the sample distribution of sample Us (Siegel, 1956:121).

8.5.7.2 Kruskal-Wallis One-way analysis of variance for independent groups

The researcher applies the Kruskal-Wallis one-way analysis of variance to determine whether κ independent samples from different populations show a significant difference. Two independent samples are required (Theron, 1992:364). The decision is probabilistic due to the fact that the aim is to determine whether sample differences represent chance variations or indicate genuine population differences (Siegel 1956:184). The Kruskal-Wallis statistic is used to test the null hypothesis (H_0) that κ comes from either the same population or from identical populations with respect to averages. It shows whether the sum of the ranks are sufficiently disparate so that the researcher can be sure that they are not likely to have been derived from samples from identical populations. Daniel (1978:201-202) offers a formula for calculating the Kruskal-Wallis statistic (H):

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{1}{n_j} \left[R_j - \frac{n_j(N+1)}{2} \right]^2$$

where R_j is the sum of the ranks assigned to observations of the j^{th} sample and n_j (N+1) / 2 is the expected sum of squares (Daniel, 1978:202).

8.5.8 NON-PARAMETRIC MEASURES OF ASSOCIATION

8.5.8.1 Nominal measures

Nominal measures of association include x^2 (chi-square), Cramer's V, Lambda (λ), Goodman and Kruskal's tau, the uncertainty coefficient, and Kappa. Only the chi-square test will be used in this study; the rest will therefore not be discussed.

Bohrnstedt et al (1988:490) view the chi-square (x^2) statistic as an appropriate test for assessing the statistical significance of crosstabulated variables. The test is based on a comparison between the (observed) cell frequencies of a crosstabulation with the frequencies that would be expected in the case where the hypothesis of no relationship was true. The values of the chi-square statistic are always positive (non-negative). This implies that the values may vary in value from zero to plus infinity ($+\infty$) (Bohrnstedt et al, 1988:121).

8.5.8.2 Ordinal measures

Non-parametric measures on ordinal level are Kendall's tau_β (t_β) and Spearman's rho (*r*) as well as Somer's d. The Spearman's rho correlation is a popular ordinal measure while Kendall's tau_β is also one of the most widely used ordinal techniques.

When continuous variables have too many abnormalities, a correction is needed. In such a case the scores are usually reduced to ranks and then calculated with Spearman's rho. Two sets of rankings (on the same two variables) are compared by:

firstly taking the difference of ranks (D_i);

then squaring the difference in ranks (D_i^2); and

lastly, adding up the squared differences:

$$\sum_{i=1}^{n} D_i^2$$

This value is placed in the formula:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} (D_i^2)}{(N)(N^2 - 1)}$$

(Bohrnstedt et al, 1988:326).

with r_s being the sample estimate of the population parameter, P_s.

8.5.9 MULTIPLE REGRESSION

In an attempt to improve on the simple linear-regression model, the accuracy of a prediction can be increased through incorporating additional information from several independent variables (Mason et al, 1989:182). This is referred to as multiple regression, and the simplest form is when the scores on two independent variables (*X1* and *X2*) are used to predict the score on *Y*. The multiple regression coefficient indicates the strength of the association between a continuous dependent variable and an independent variable while controlling the other independent variable in the equation (Bohrnstedt et al, 1988: 495-496).

Cooper et al (1998) state that multiple regression can be used as a descriptive tool in various types of situations:

- When developing a self-weighting estimating equation to predict values for a criterion variable (DV).
- It can be a descriptive application. This calls for controlling of confounding variables to better evaluate the contribution of other variables.
- It can also be used to test for and explain causal theories (referred to as path analysis). Here multiple regression is used to describe the linkages that have been advanced from a causal theory.

The regression coefficient may be stated either in raw score units or as a standardised coefficient (Cooper et al, 1998:563). In both these cases the coefficient value states the amount that Y varies for each unit change of the associated X variable, while the effects of all other X variables are being held constant (Cooper et al, 1998:563).

8.6 CONCLUSION

In this chapter the research methodology and design was discussed. The research strategy was explained, after which the process of survey research was discussed in detail by referring to the objectives of the study. The population was demarcated and the procedures for administering the questionnaires and the collection of data were discussed. The last part of the chapter entailed a description of the statistical methods to be used, viz descriptive statistics, analysis of variance, discriminant analysis, correlation statistics, Student's T-test, non-parametric statistics, and multiple regression. A description of the sample characteristics will follow in Chapter 9.